Doctoral Thesis

Data Assimilation of Remote Sensing Soil Moisture Retrievals with Local Ensemble Transform Kalman Filter Scheme

Eunkyo Seo

Department of Urban and Environmental Engineering
(Environmental Science and Engineering)

Graduate School of UNIST

2019

Data Assimilation of Remote Sensing Soil Moisture Retrievals with Local Ensemble Transform Kalman Filter scheme

Eunkyo Seo

Department of Urban and Environmental Engineering
(Environmental Science and Engineering)

Graduate School of UNIST

Data Assimilation of Remote Sensing Soil Moisture Retrievals with Local Ensemble Transform Kalman Filter scheme

A thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Eunkyo Seo

June 4, 2019
Approved by

Advisor
Myong-In Lee

Data Assimilation of Remote Sensing Soil Moisture Retrievals with Local Ensemble Transform Kalman Filter scheme

Eunkyo Seo

This certifies that the thesis/dissertation of Eunkyo Seo is approved.

06/04/2019

signature

Advisor: Myong-In Lee

signature

Thesis Committee Member: Dong-Hyun Cha

signature

Thesis Committee Member: Jungho Im

signature

Thesis Committee Member: Jee-Hoon Jeong

signature

Thesis Committee Member: Yoo-Geun Ham

Abstract

This thesis proposes a development of land data assimilation system to produce realistic land surface states, which is performed with diverse remote sensing retrievals using advanced land surface model (LSM) and data assimilation techniques. Remote sensed soil moisture retrievals with high-temporal and –spatial resolution is recently available. For instance, the Advanced Scatterometer (ASCAT), a C-band active microwave remote sensing instrument, and the NASA Soil Moisture Active Passive (SMAP) L-band passive remote sensing retrievals provide global near-surface soil moisture condition in real-time. Bias corrected observation datasets are used in the assimilation based on cumulative distribution function fitting because there is a large discrepancy of soil moisture contents between each retrieval and LSM offline simulation by difference sensed layer depth and algorithms, and characteristics of model physics. This study performs the soil moisture data assimilation using these bias corrected satellite retrievals with Local Ensemble Transform Kalman Filter (LETKF) scheme. The impact of the soil moisture assimilation is evaluated with ground based in-situ soil moisture measurement network over the globe.

The result reveals that each satellite retrieval provides significant added value in the data assimilation. The impact of the assimilation tends to be better improved when active and passive satellite retrievals are simultaneously used. The temporal correlation of the assimilated soil moisture increases surface soil moisture skills by $\Delta R \approx 0.12$ over the continental U.S., and the improvement at root-zone is $\Delta R \approx 0.1$. The result is explained by "Assimilation Gain", where the quality of assimilated satellite data and the number of assimilated observations strongly contribute to the skill improvement of assimilated soil moisture estimates. The skill improvement through the multi-retrieval assimilation is mostly significant in transitional climate regime where land-atmosphere interaction is strong, and the impact of soil moisture initialization is clearly shown in the forecast model. Furthermore, the skill improvement is also significant in other validated regions (e.g. western Europe, and central Tibetan Plateau). The magnitude of the skill improvement through the assimilation is large when the quality of satellite retrievals tends to be better than that of the open loop. The assimilated soil moisture estimates are widely used in understanding land surface physical process of hydrology cycle and the land-atmosphere interaction. Furthermore, the realistic land surface condition gives the better information in land surface monitoring system.

Contents

List of figures

List of tables

Chapter 1

1. Introduction

1.1. Background and motivation

The hydrological and thermodynamic land surface conditions affect and are affected by several agricultural and hydrological processes, drought development, runoff generation and many other processes. The qualified high-resolution estimates of land energy and water states are profoundly significant for understanding its impacts on the regional and global climate system through atmospheric since the regional change of land surface conditions through the exchange of moisture and energy budgets in land–atmosphere interaction critically influences on the forming the remote response as well as the local land surface process. For instance, the imposed dryness in the North America could excite a robust circumglobal teleconnection response across the Northern Hemisphere midlatitudes in the boreal summer season (Teng, Branstator, Tawfik, & Callaghan, 2019), where nonlinear processes associated with synoptic eddies at jet exit regions play an important role in forming remote response from North America and North Atlantic to Eurasia and the North Pacific Ocean.

The importance of the land surface boundary conditions in dynamical forecast systems has been also increased in subseasonal time scale (approximately 2 weeks to about 2 months) prediction such as land surface can also contribute to low-frequency land surface variability (Conil, Douville, & Tyteca, 2007; Dirmeyer, 2005; Herve Douville & Chauvin, 2000; Ford, Dirmeyer, & Benson, 2018). The accurate characterization of the land surface can improve weather and climate prediction in dynamical forecast models. Recently, numerous studies have addressed the increase of forecast skill and/or potential predictability of atmosphereic conditions through improving land surface process represented by dynamical forecast models (Dirmeyer, 2003; H Douville, 2004; Guo, Dirmeyer, & DelSole, 2011; R. Koster et al., 2011; R. D. Koster, Dirmeyer, et al., 2004; R. D. Koster et al., 2002; R. D. Koster et al., 2010; R. D. Koster, Suarez, et al., 2004; R. D. Koster et al., 2006; Materia et al., 2014; Prodhomme, Doblas-Reyes, Bellprat, & Dutra, 2016; Seneviratne, Koster, et al., 2006; Seo et al., 2018). The realisitc land-atmosphere interaction through the soil moisture initialization also leads to the better represention of boreal summer extreme climate events (e.g. heatwave or drought) in the forecast models. For instance, the land surface initializaiton with realistic soil moisture significantly improves the simulatuion for record-breaking 2003 and 2010 European heatwave events (Ferranti & Viterbo, 2006; Fischer, Seneviratne, Vidale, Lüthi, & Schär, 2007; Hauser, Orth, & Seneviratne, 2016; D. G. Miralles, Teuling,

Van Heerwaarden, & de Arellano, 2014; Orth, Dutra, & Pappenberger, 2016; Quesada, Vautard, Yiou, Hirschi, & Seneviratne, 2012; Seneviratne et al., 2010; Seneviratne, Lüthi, Litschi, & Schär, 2006; Bart van den Hurk et al., 2012; Weisheimer, Doblas-Reyes, Jung, & Palmer, 2011; Zaitchik, Macalady, Bonneau, & Smith, 2006). Furthermore, the influence of the snow cover on anoalous atmospheric teleconnections during boreal winter–spring season has received attention for its impact on seasonal prediction (Jeong et al., 2013; Orsolini et al., 2013). These comprehensive understanding land–atmosphere interaction in dynamical forecast systems achieves the increasing forecast skill not only for boreal summer season but also for the winter season. For future climate projection, as a driver of boundary layer properties, the role of the land surface is increased with the strengthen land–atmosphere coupling and the expansion of the coupling area, particularly toward cooler and wetter regions, which implies that some trends for severe droughts and heat waves may be manifested (Dirmeyer, Jin, Singh, & Yan, 2013). Therefore, the understanding of water and energy cyle at land surface layer with mostly realistic land surface states is important not only for the local change of land–atmosphere feedback but also for the increased impact of circumglobal stationary wave events in the recent decade boreal summers.

Especially, soil moisture plays an important variable in the land-atmosphere interaction, since it controls the energy and water balance at land surface interface and stores them in sub-surface layer. Soil moisture information can be obtained via three main ways: (1) ground based in situ measurements, (2) remote sensing retrievals from low-frequency active and passive microwave sensors, and (3) simulation of LSM driven by meteorological variables. However, they have distinctive disadvantages. In situ measurements provide most reliable land information of surface and sub subsurface layers at a station but they have limitations in terms of spatial and temporal resolution, although the quality of the observation is mostly reliable. Remotely sensed soil moisture provides mostly surface information up to ~5 cm penetration depth and is affected by errors due to spectrotemporal coverage, the heterogeneity of land cover, and retrieval algorithm even though most applicational fields also require root zone information. LSM offline integration contains large uncertainties in the model physical parameter and the meteorological forcing variables. In order to overcome addressed problem in each soil moisture observation, land data assimilation system is a state-of-the-art to combine the complementary soil moisture information from diverse observations (Reichle & Koster, 2005).

Based on these available soil moisture information, qualified soil moisture estimates are mainly determined by model physical process, the quality of the near-surface atmospheric boundary forcing variables driving LSM, and data assimilation methodology (Kumar et al., 2006; Kumar et al., 2008; Sheffield, Goteti, & Wood, 2006). In previous studies, the importance of physical parameterizations in LSM and boundary conditions has been addressed. For instance, the Global Land Data Assimilation

System (GLDAS) ingests satellite- and ground-based observational datasets with advanced LSM to obtain long-term optimal land surface states and fluxes (Rodell et al., 2004). A trial to produce the realistic land surface hydrological fields have been done by correcting the precipitation forcing with observational surface meteorological forcing variables and revised parameter values in the LSM (Reichle et al., 2011). Moreover, many of previous studies evaluate the assimilation of land surface conditions with abundant land datasets to constraining model errors and considering forecast and observational errors (Andreadis & Lettenmaier, 2006; Bosilovich, Radakovich, da SILVA, Todling, & Verter, 2007; Crow & Wood, 2003; Drusch, 2007; Margulis, McLaughlin, Entekhabi, & Dunne, 2002; Reichle & Koster, 2005; Rodell & Houser, 2004; BJJM Van Den Hurk, Jia, Jacobs, Menenti, & Li, 2002). However, there are not many studies that tries to examine the impact of data assimilation scheme and adopted remote sensing soil moisture retrievals, relatively.

This study performs the soil moisture data assimilation using recently available remote sensed soil moisture retrievals with Local Ensemble Transform Kalman Filter (LETKF) scheme. Based on this baseline data assimilation framework, this study examines several experiments with different assimilation schemes and satellite datasets to optimize the data assimilation system. The realistic soil moisture states produced by the data assimilation system is widely used in various fields related to understanding land surface processes.

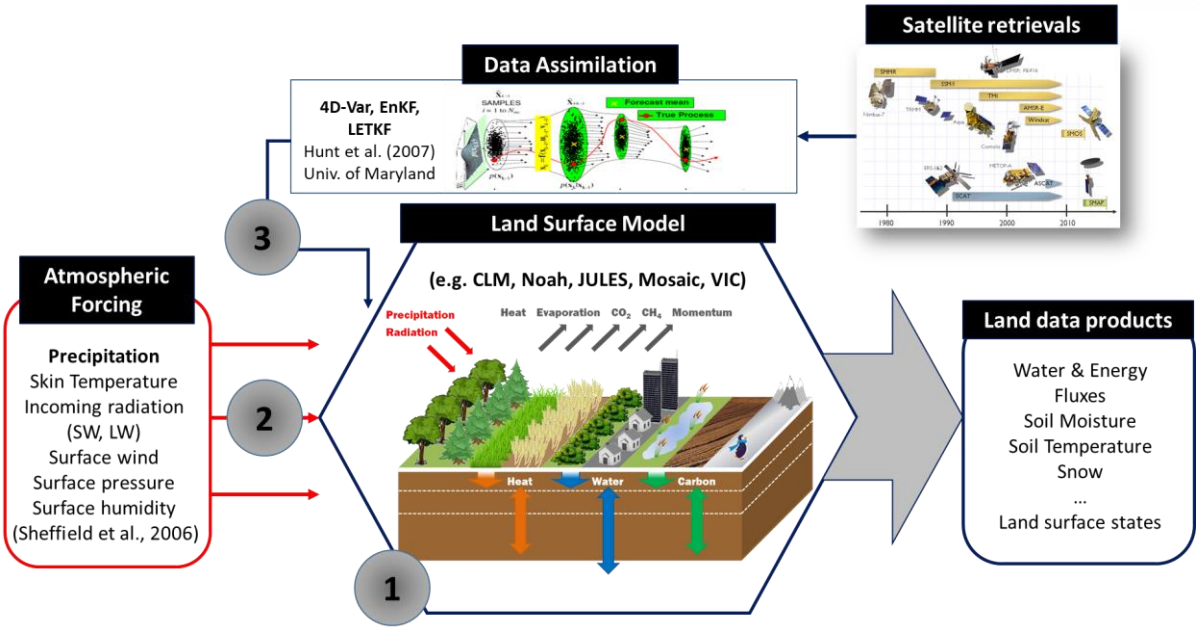

Figure 1.1 Schematic of production of land surface variables in data assimilation system.

1.2. Objectives

This study develops global land data assimilation system with LETKF scheme using remote sensing observations and evaluates the estimates of assimilated soil moisture comparing with ground based in-situ measurement network over the globe. Based on this perspective, this study aims to understand LETKF scheme and optimize it into newly developed land data assimilation system and evaluates the skill gained by assimilation of ASCAT, SMAP, and both.

Chapter 2 describes remotely sensed soil moisture datasets and ground-based in situ measurements, the information of LSM and assimilation schemes employed in this study and used evaluation metrics. Chapter 3 represents the framework of land data assimilation system and diagnoses the land surface variables produced by the assimilation system. Conclusion is given in chapter 4. Chapter 5 provides outlook and future works.

Chapter 2

2. Establishment of land data assimilation system

2.1. Background

As increasing the interest of understanding land surface process with realistic land surface information, there are several previous studies to establish land data assimilation framework (Draper, Reichle, De Lannoy, & Liu, 2012; Kolassa, Reichle, & Draper, 2017; Reichle, 2008; Reichle & Koster, 2005; Reichle et al., 2007; Reichle, Walker, Koster, & Houser, 2002). Most of previous studies for the land data assimilation has adopted Ensemble Kalman Filter (EnKF) scheme (Evensen, 2003, 2009; P. Houtekamer & Zhang, 2016), which is one of most representative non-variational method. Background error covariance in the non-variational method is sequentially obtained from background ensemble perturbation for each time step, which can consider the nonlinearity of the background states. The non-variational approach is appropriate in the representation of flexible error in model dynamical parameters and is feasible for the modestly nonlinear and intermittent character of land surface conditions (Reichle, McLaughlin, & Entekhabi, 2002). However, the scheme has well-known problems: (1) sampling issue (Impossible to resolve degree of freedom of dynamical model due to limitation of ensemble members), (2) covariance localization (demanding computing resource to calculate large size matrix), and (3) not enough ensemble spread. These problems are practically improved by localization and transformation. For large modeling domains or high spatial resolution of the land data assimilation, the EnKF state matrix with a large number of variables may demand an expansive computation. Moreover, the practical limitation of ensemble members leads to spurious sample correlation a large distance point. Because the length scale in land surface variables is typically shorter than those in atmospheric states, the EnKF requires a suitable localization for the error covariances. The localization is efficient in the requirement of computing resource due to only calculating localized error covariance rather than large domain. The transformation includes analysis ensemble perturbation transformed into the analysis field (square root filter), which can control analysis ensemble spread through modifying inflation parameter. Therefore, this study adopts a recently advanced assimilation method, LETKF (Hunt, Kostelich, & Szunyogh, 2007; Miyoshi & Yamane, 2007), to operate land data assimilation system.

For the soil moisture observation, in recent decade, useful remote sensed near-surface soil moisture retrievals with high–temporal and –spatial resolution are also available. For instance, the passive microwave Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E)

provided valuable soil moisture observations between June 2002 and October 2011 and was replaced by the Advanced Microwave Scanning Radiometer 2 (AMSR2) from May 2012 to current days. Compared to AMSR-E, AMSR2 has a larger passive microwave antenna to achieve better spatial resolution and a C-band frequency (7.3 GHz) channel in addition to the original C-band frequency (6.9 GHz) to mitigate radio-frequency interference (RFI) effects. However, the improvements in AMSR2 have not led to statistically significant differences in performance for Land Parameter Retrieval Model (LPRM) retrievals, when compared with AMSR-E (Cho, Su, Ryu, Kim, & Choi, 2017). The Advanced Scatterometer (ASCAT), an active microwave remote sensing instrument, provides near real-time soil moisture information since October 2006 (Wagner et al., 2013). Compared to AMSR-E, ASCAT tends to better representation of relative soil moisture variation across Europe, although they are commonly C-band microwave remote sensing instruments (Brocca et al., 2011). NASA Soil Moisture Active Passive (SMAP) L-band active (radar) and passive (radiometer) microwave remote sensed retrieval also provides global mapping of high-resolution soil moisture since April 2015, but active sensor will not be possible due to the sudden failure of SMAP's radar in July 2015 (Chan et al., 2016). The Soil Moisture Climate Change Initiative (CCI) project is an ESA Program of Global Monitoring of Essential Climate Variables (ECV). The project has tried to create a temporally and spatially high-resolution consistent soil moisture time series with collaborating diverse satellite measurements, based on active and passive data (Wouter Dorigo et al., 2017; Gruber, Dorigo, Crow, & Wagner, 2017; Y. Liu et al., 2012).

Based on these data assimilation schemes and remote sensing soil moisture datasets, this study has a purpose to develop an optimized land surface data assimilation system. All of datasets used in this study are obtained in near real-time which is feasible to implement on the operational forecast systems. In this section, we will describe the utilized assimilation schemes and observational datasets.

2.2. Land surface model

2.2.1. The Joint UK Land Environment Simulator (JULES)

This study mainly uses the energy and water components of a new community land surface model called the Joint UK Land Environment Simulator (JULES) (Best et al., 2011). The model is developed by the U.K. Met Office and originated from the Met Office Surface Exchange Scheme (MOSES). It has been coupled to an atmospheric component and operated to weather and climate prediction systems. Based on water and energy balance equations, JULES prognosticates the land variables from the surface to the underground 3 meters below using 4 layers (0.1, 0.25, 0.65, and 2 meters). Land cover in the JULES consists of five plant functional (Broadleaf trees, Needleleaf trees, C3 (temperate) grass, C4

(tropical) grass, and Shrubs) and four non-vegetation types (Urban, Inland water, Bare soil, and Land-ice), respectively. Surface related variables (e.g. albedo, roughness length, and so on) are specified by the pre-determined land cover. After precising the surface variables depending on the land cover, the initialization of the prognostic variables within the JULES model can integrate the land surface variables driven by the meteorological forcing variables. In the data assimilation process, we update a prognostic value associated with the soil moisture (Moisture concentration of each soil layer, which is contained in model prognostic value) for every assimilation cycle. The JULES integration is carried out at 50 km spatial resolution and 4 layers vertical resolution.

2.2. Data

2.2.1. Atmospheric boundary conditions

For producing qualified land surface variables in off-line LSM calculation, atmospheric boundary condition is another important contributor, which includes 2-m air temperature and humidity, precipitation, 10-m wind speed, radiative fluxes, and pressure at surface. These variables are obtained from the 6-hourly Japanese 55-year Reanalysis (JRA55) (Kobayashi et al., 2015). Precipitation, which is a most critical input for determining soil moisture accuracy in the land surface modeling, has been primarily corrected by the monthly Climate Prediction Center Merged Analysis of Precipitation (CMAP) (Xie & Arkin, 1997) and the Global Satellite Mapping of Precipitation (GSMaP) (Aonashi et al., 2009; Kubota et al., 2007; Ushio et al., 2003; Ushio et al., 2009). The GSMaP project was promoted by the JAXA Precipitation Measuring Mission (PMM) and it provides hourly gauge-calibrated rain rate with 10 km spatial resolution over the global domain (60°S–60°N). Therefore, the precipitation boundary forcing is replaced with the 6-hourly averaged GSMaP rainfall data (60°S–60°N) and the monthly mean rainfall of the reanalysis is corrected by the monthly CMAP observation over the residual region where the GSMaP does not cover (90°S–60°S and 60°N–90°N).

2.2.2. In situ soil moisture measurements

The International Soil Moisture Network (ISMN; http://www.ipf.tuwien.ac.at/insitu) collects available in situ soil moisture measurements from operational networks and validation campaigns over the globe (WA Dorigo et al., 2011). The ISMN is an international cooperation to gather a global in situ soil moisture measurement. It is ground-based soil moisture direct measurement from surface to subsurface levels. Figure 2.1 shows the global distribution of qualified ISMN station points. The density

of soil moisture network is different by each continent. For instance, most of soil moisture data are concentrated in North America and some of them are exist in Eurasia continent. The interannual variation of available soil moisture data has been increased after 1980. However, the increment of the observation is only focused on North America, but the tendency is not consistent on Eurasia continent (Fig. 2.2). In-situ stations in east Asia is mainly available before 2000s and soil moisture data over Europe is only available fewer than 500 and 100 for all land surface and surface level, respectively.

The data provides hourly soil moisture measuring the dielectric constant of the soil and the sensors are vertically inserted at depths of 5, 10, 20, and 100 cm. It also provides automatic detection of problematic observations as flagged with data quality, and we only use "Good" qualified soil moisture information and discard obviously unrealistic value such as data beyond the physically possible range. Furthermore, we also exclude the frozen state of soil moisture in which the temperature information is based on soil temperature measurement. After the quality control process with the hourly data, we recalculate daily mean soil moisture, and then additionally screen for sufficient data availability at least 90 daily observation samples during the research period. The qualified in situ datasets are used for the validation of surface and subsurface layers, where the surface and root-zone soil moistures are the measured value at 5 cm depth and a depth-weighted average of in situ sensors at 5, 10 and 20 cm.

To validate in situ observations, we evaluate skill in terms of the temporal correlation coefficient of daily averaged anomalies at in situ measurement sites compared with ASCAT and SMAP soil moisture retrievals and the non-assimilated ensemble LSM simulation (hereafter referred to as the "open loop"). Based on these skills at in situ observation, the quality of in situ sites is controlled, where either ASCAT and SMAP retrieval skill is substantially worse than the corresponding open loop skill ($R_{retrieval} - R_{openloop} < -0.2$) are excluded from the analysis. The selected in situ sites are distributed over the North America (Soil Climate Analysis Network, U.S. Climate Reference Network, and Snow Telemetry) and few of western Europe (Soil Moisture Observing System - Meteorological Automatic Network Integrated Application) and central Tibetan Plateau (Soil Moisture and Temperature Monitoring Network on the central Tibetan Plateau) (Fig. 2.3). They are used to evaluate the assimilated soil moisture estimates, but most of observations are concentrated in the continental United States.

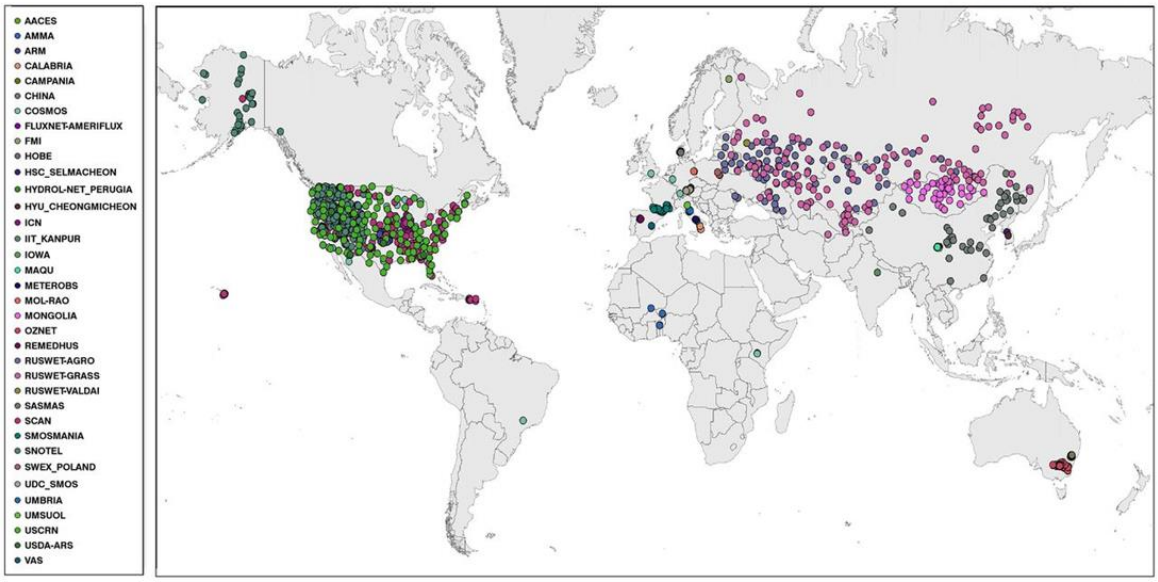

Figure 2.1 Global distribution of International Soil Moisture Network stations.

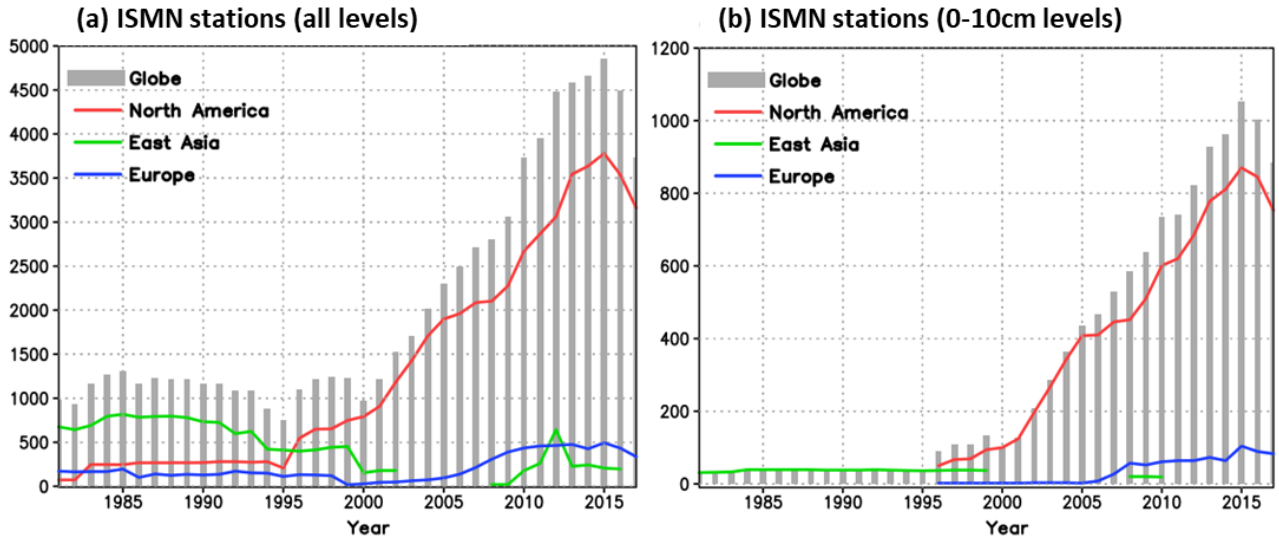

Figure 2.2 Interannual variation of available International Soil Moisture Network stations for (a) all land surface levels and (b) 0–10 cm levels.

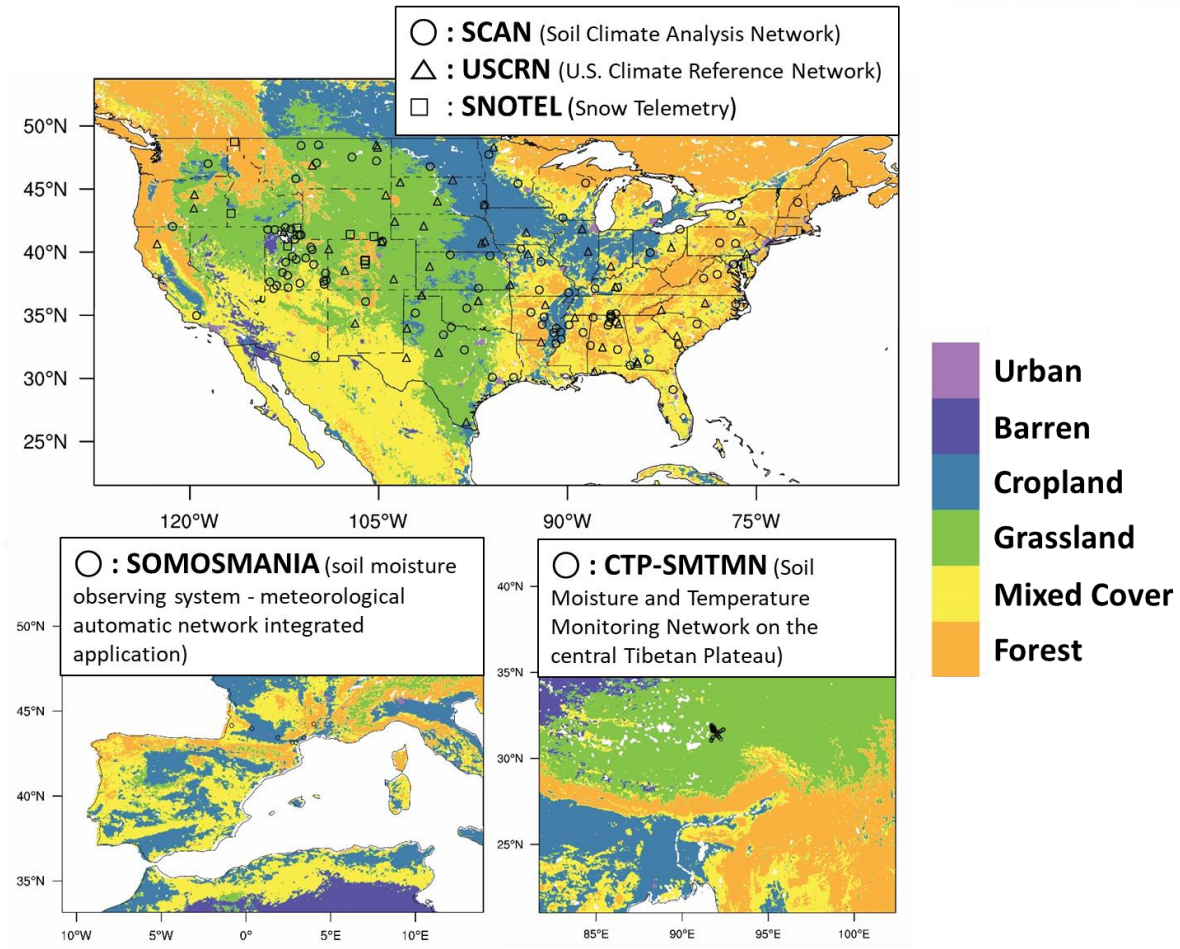

Figure 2.3 The location of (a) SCAN/USCRN/SNOTEL over the North America, (b) SOMOSMANIA in western Europe and (c) CTP-SMTMN monitoring sites. In situ measurements are used to evaluate soil moisture estimates from satellite and assimilated product, overlaid with MODIS land cover classes.

2.2.3. Remote sensing soil moisture retrievals

This study adopts real-time available global near-surface soil moisture datasets provided by ASCAT and SMAP remotely sensed retrievals. ASCAT is a C-band active (radar) microwave remote sensor, and SMAP is a L-band passive (radiometer) remote sensing retrieval. Depending on the frequency of each microwave band, the penetrated depth of the soil moisture is different, where ASCAT and SMAP are 1-2 cm and 5 cm, respectively. Radiometer has a higher sensitivity to soil dielectric properties and other surface parameters, such as surface temperature or vegetation water content, and soil moisture than that of radar in the microwave spectrum range (Paloscia & Pampaloni, 1988; Schmugge, O'Neill, & Wang, 1986). In contrast, the radar observation has a less sensitivity to surface soil variables such as wetness and temperature and the active sensor would receive multiply scattered noises over topographically complex, wetland, and forest regimes (Dobson & Ulaby, 1986). However, the spatial resolution of the radar observation had much finer resolution than that of radiometer due to the characteristics of remote sensing measurement, so that ASCAT soil moisture dataset has ~25 km resolution at frequencies used as soil moisture retrievals even though SMAP spatial resolution is about ~36 km.

Prior to the soil moisture assimilation, the quality control for the assimilated soil moisture retrievals is conducted by the characteristics of each microwave observation. Firstly, we discard all of retrieval datasets flagged by non-qualified which is provided by each satellite dataset and the range of soil moisture condition is not realistic. Additionally, all of observations are discarded in which MODIS land cover indicates forests (> 60% trees and woody vegetation) and remove grid cells with a wetland cover (indicated by ASCAT data) above 10% and ASCAT data is also discarded where topographical complex and wetland fraction are larger than 10%. The quality of ASCAT data is significantly degraded in which topographical complex and wetland cover are flagged to large percentage compared with in situ measurements. Some previous studies diagnosed the quality of soil moisture from remote sensing retrievals compared to in situ measurements. ASCAT represents that the daily temporal correlation skill over the North America is approximately 0.5 compared to ground base in situ observation (Albergel et al., 2012) and SMAP shows the correlation is about 0.7 (Pan, Cai, Chaney, Entekhabi, & Wood, 2016). It means that the skill of SMAP soil moisture observation is better than that of ASCAT. For the observation errors, a spatially and temporally constant error standard deviation of 10% is used in the ASCAT (Wouter Dorigo et al., 2010) and the standard deviation of SMAP errors is 0.04 m^3 m^{-3} (Chan et al., 2016). When the bias of the soil moisture retrievals is corrected based on CDF filtering (discussed in the Section 2.3), the specified standard deviations of the observation error are also rescaled with the same method at each grid cell.

2.2.4. MODIS land cover

The information of land cover is adopted from MODIS Collection 5 global land cover (Friedl et al., 2010). The most important improvement in the updated version is to produce global land cover dataset with 500 m spatial resolution. Compared to previous version, the input data and classification algorithm used in the production of the Collection 5 MODIS Global Land Cover Type is also substantially revised. For instance, land surface temperature is included as an input variable in the revised training site, and ancillary datasets is refined in the post-processing of ensemble decision tree results. The Collection 5 consists of the 17-class International Geosphere–Biosphere Programme classification (IGBP) (Loveland & Belward, 1997); Evergreen needleleaf forests (1), evergreen broadleaf forests (2), deciduous broadleaf forests (3), deciduous needleleaf forests (4), mixed forests (5), closed shrubland (6), open shrublands (7), woody savannas (8), savannas (9), grasslands (10), permanent wetlands (11), croplands (12), urban and built-up lands (13), cropland/natural vegetation mosaics (14), permanent snow and ice (15), barren (16), and water (17). Based on these 17 classed land covers, forests, mixed land cover, grassland, and cropland are defined all of forests flagged land cover (1-5), shrublands, woodland and vegetation mosaics (6-9 and 14), grasslands (10) and croplands (12), respectively. In this study, we categorize the skill improvement of assimilated estimates depending on classified land covers.

2.3 Bias correction

There is large discrepancy of soil moisture contents between remote sensing retrievals and LSM integrations of observed meteorological forcing variables although surface soil moisture data seems to be consistent and useful information. This bias is possibly induced by difference sensed layer depth and characteristics of model physics. For instance, the mean climatology of soil moisture from the satellite and model simulation shows a large discrepancy. The JULES offline simulation shows wet bias compared to the ASCAT satellite data over the South and North America, western Europe, and the Australia and there is dry bias over the mid-latitude Eurasia continent (Fig. 2.4). A method to correct the addressed bias for the soil moisture is to match the cumulative distribution function (CDF) fitting of the satellite and model simulation (Reichle & Koster, 2004). In the North America, the CDF of soil moisture concentration of ASCAT, SMAP, and the JULES offline simulation is significantly different for each other (Fig. 2.5), so that the raw data of ASCAT and SMAP soil moisture satellites is rescaled to wetter soil moisture value corresponding to the LSM climatology. Therefore, in this study, bias corrected observation datasets are used in the assimilation based on CDF fitting.

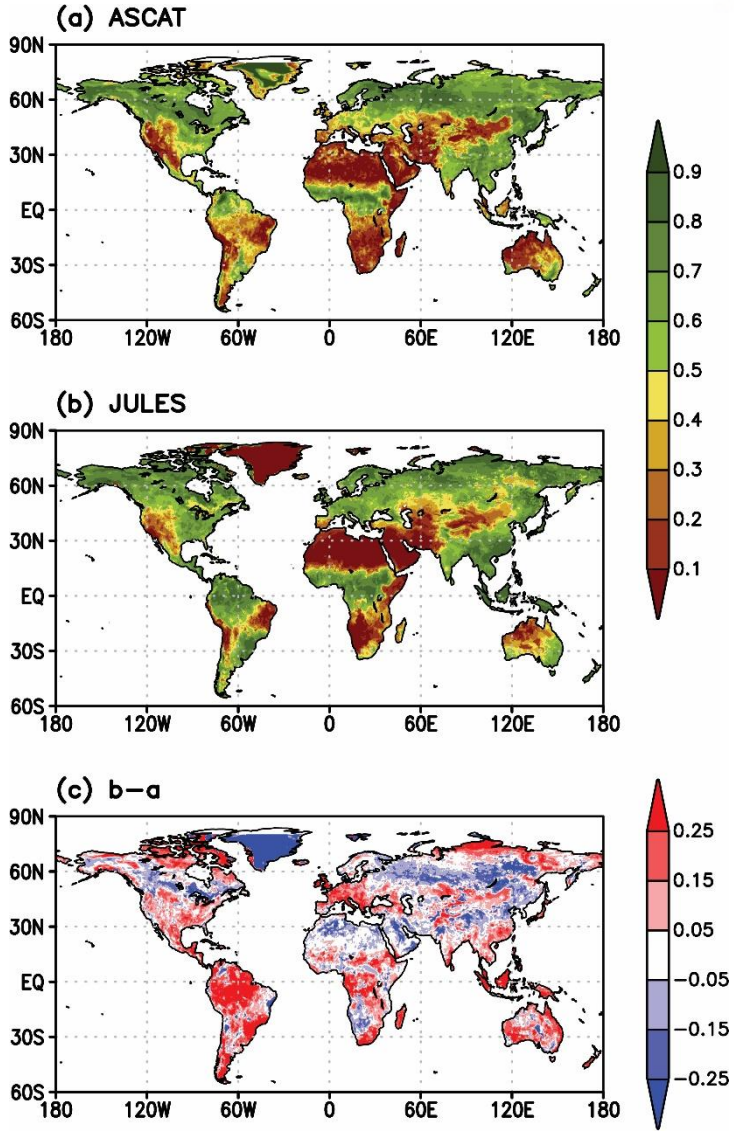

Figure 2.4 The climatology of soil moisture concentration during May–September 2016 from (a) ASCAT satellite, (b) the JULES offline calculation and (c) the difference of b minus a.

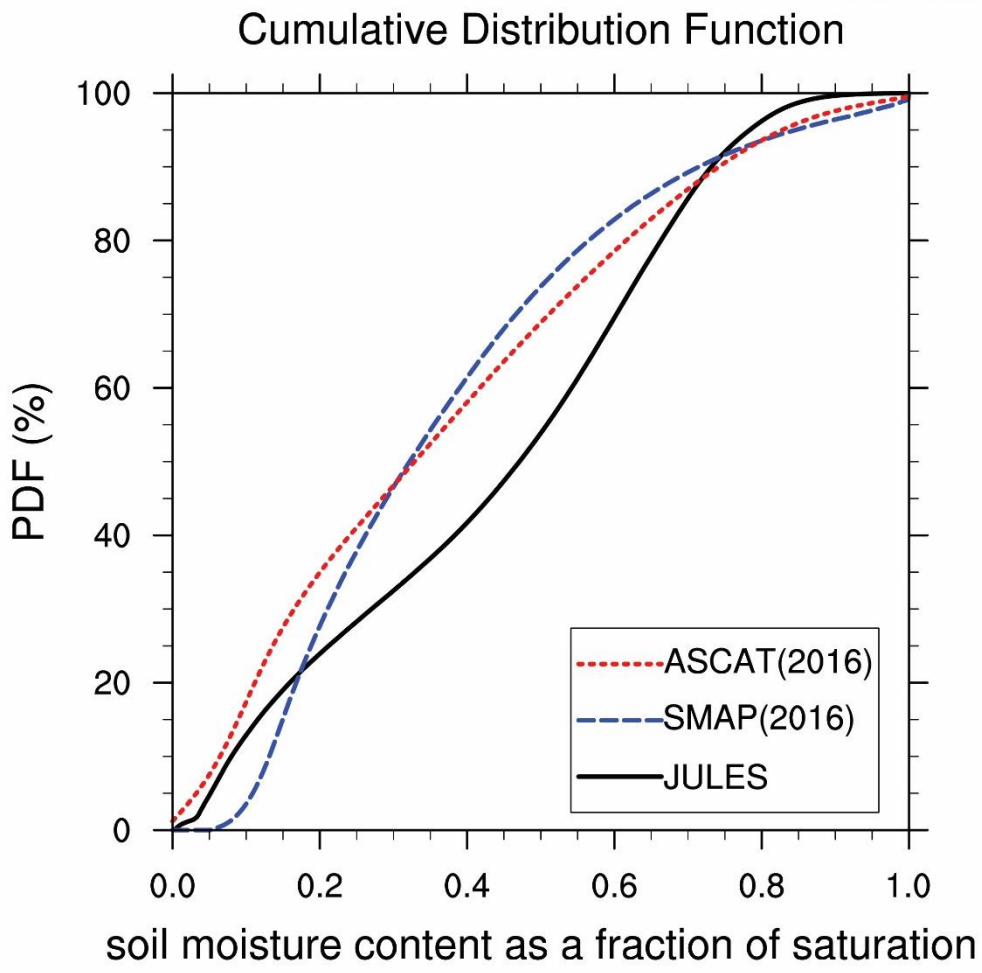

Figure 2.5 The cumulative distribution function of soil moisture concentration of ASCAT, SMAP, and the JULES offline simulation over the North America (130°W–90°W, 20°N–50°N) during May–September 2016.

2.4. Data assimilation methods

2.4.1. Ensemble Kalman Filter (EnKF)

The ensemble Kalman Filter (EnKF) is originated from Kalman filter theory and Monte Carlo estimation method and it is a well-known non-variational data assimilation technique (Evensen, 1994). After introduced the basic concept of EnKF, many other studies have tried to improve the original algorithm. The overall review of the EnKF scheme is addressed in P. Houtekamer and Zhang (2016). EnKF scheme has strong and weak points and its feature is given in Table 2.1. This scheme has positive aspect in terms of the automatic estimation of flow-dependent background error covariance. By this advantage, non-variational method is feasible in climate shift such as seasonal variation and climate change, because the model itself can estimate the background error. Furthermore, it is relatively easy application to many kinds of dynamical models compared to variational method and ensemble model forecast can estimate the nonlinearity of the background states. In contrast, it has also some disadvantage in terms of sampling issue, covariance localization and ensemble spread. The ensemble sampling issue is a chronic problem of the non-variational assimilation method since ensemble members are limited to resolve degree of freedom of dynamical model. The problem of covariance localization and ensemble spread is addressed for the limitation of computing resource due to calculating large size matrix and the underestimated background error, respectively.

Table 2.1 Introduction of strong and weak points of EnKF scheme

Strong points	Weak points
1. Automatic estimation of flow-dependent background error covariance - Important in the condition of climate shift (e.g. global warming, seasonal variation)	**1. Sampling problem** - Impossible to resolve degree of freedom of dynamical model due to limitation of ensemble members
2. Ease of implementation - Relatively easy application to many kinds of dynamical models	**2. Covariance localization** - Limitation of computing resource due to calculating large size matrix
3. Ensemble forecast - Considering nonlinearity of background states	**3. Ensemble spread** - Innovation is guaranteed under enough background error covariance

The implementation of EnKF updates a prior background estimate $X_b(t)$ of soil moisture prognostic value with an observation y_0 at a time t to an updated analysis $X_a(t)$. Base on this formulation, the innovation of the observation is determined by the weighting of background and observation error covariance. The background error is reflected by background ensemble perturbation for every assimilation cycle in the non-variational assimilation scheme, so that the scheme can calculate automatic estimation of flow-dependent background error covariance. Kalman gain matrix K gives appropriate weight to observational increments and it is ratio of the background error (P_b) compared to the sum of background and observation error (R). Forward operator $\mathcal{H}$ is used for mapping from model to observation space. The perturbation of the observation explains measurement instrument errors and representative errors, which adds the gaussian distributed random numbers v_0 with number of ensembles N where the mean of random numbers is zero and its variance is determined by the observation error covariance R. After calculating soil moisture analysis, the updated prognostic fields are projected onto next time step background forecasts with the nonlinear operator $\mathcal{M}[\cdot]$ including all deterministic forcing data value. Additionally, uncertainties related to errors in LSM physics or atmospheric boundary forcing variables in the model error w_b which assumes zero mean random variables with model error covariances.

$$X_a(t) = X_b(t) + K[y_0 + v_0 - \mathcal{H}X_b(t)] \tag{2.1}$$

$$K = P_b\mathcal{H}^T(\mathcal{H}P_b\mathcal{H}^T + R)^{-1} \tag{2.2}$$

$$P_b = \frac{1}{N-1}\sum_{i=1}^{N}(x_b^i - \overline{x_b})(x_b^i - \overline{x_b})^T \tag{2.3}$$

$$X_b(t+1) = \mathcal{M}[X_a(t)] + w_b \tag{2.4}$$

The understanding matrix size for each component on the introduced equations is important to calculate the analysis field. The shape of $X_a(t)$ and $X_b(t)$ is $m \times N$ matrix composed of model grids m and ensemble number N. Correspondingly, the shape of $y_0 + v_0$ and $\mathcal{H}X_b(t)$ is $n \times N$ matrix composed of the number of observations n and ensemble numbers. The role of the Kalman gain is weighting the difference of observation and model background and coverts the matrix shape from the observation to the model, so that its shape is $m \times n$.

2.4.2. Local Ensemble Transform Kalman Filter (LETKF)

LETKF was initially proposed by Hunt et al. (2007) which is an advanced version of Local

Ensemble Kalman Filter (LEKF) (Ott et al., 2004; Ott et al., 2002). Even though there is insignificant difference in convergence between LETKF and LEKF, LETKF is the better option since LETKF is quietly faster than LEKF. A dominant advantage of LETKF scheme is the efficiency of parallel computation since the assimilation method divides the entire basin grid into independent local patches. The repetition of data assimilation within each independent local patch demands less computational memory.

Table 2.2 Introduction of strong and weak points of LETKF scheme

Strong points	Weak points
1. Automatic estimation of flow-dependent background error covariance - Important in the condition of climate shift (e.g. global warming, seasonal variation)	**1. Sampling problem** - Impossible to resolve degree of freedom of dynamical model due to limitation of ensemble members
2. Ease of implementation - Relatively easy application to many kinds of dynamical models	**2. Covariance localization (LETKF)** - Decomposing an entire global analysis domain into a number of independent sub-domains - Reducing computation (easy to parallelize feasible in high-resolution LSM and observations)
3. Ensemble forecast - Considering nonlinearity of background states	**3. Ensemble spread (LETKF)** - Including analysis error covariance where the inflation parameter is able to control the spread

Table 2.3 Comparison of data assimilation schemes by non-variational (LETKF) and variational (4D-Var) method

	LETKF	**4D-Var**
"advanced" method?	Yes	Yes
Simple to code?	Yes	No (Minimize Cost function)
Model dependency?	No	Yes
Background error covariance	Only forward	Adjoint required
Analysis errors?	Yes (ensemble perturb)	No
Limitation	Ensemble sizes	Assimilation window

Table 2.2 displays strong and weak points of LETKF scheme, and the main differences compared to the EnKF is highlighted with red color. Through the application of localization and transform on EnKF, the LETKF quietly overcomes the computing resource issue and the underestimating background fields although Sampling problem is same as EnKF. The localization decomposes the global analysis domain into several independent sub-domains, where the demanding of computation is reduced, and it is easy to parallelize feasible in high-resolution LSM and observations. In addition, the transform is enabled to transform subsequently the ensemble of observation and background errors into an analysis ensemble perturbation based on the square root of analysis error covariance matrix.

Table 2.3 is the comparison of non-variational (LETKF) and variational (4D-Var) data assimilation scheme. LETKF and 4D-Var are the most advanced data assimilation scheme among non-variational and variational method, respectively. Both of methods are different approach to compute observational innovation. In the non-variational method, the innovation is determined by the ratio of background error covariance compared to total ensemble spread, but the variational method tries to minimize, iteratively, cost function which represents model background and observational error. However, the iteration process to minimize the cost function demands lots of computation compared to simply calculation of the analysis error covariance. Furthermore, the non-variational scheme has another profit in terms of applicableness since the method can calculate the model background error covariance with their ensemble perturbations, but the variational method needs the adjoint function to represent the model error. The perturbed analysis states contain the analysis error which could resolve the parts that are not explained in nonlinear model, but variational scheme does not contain the analysis error. However, the non-variational method also has a representative problem for the limited ensemble member generation. Ideally, the number of ensemble member should be same as degree of freedom of the background model, but it is a realistic impossibility.

The LETKF scheme firstly separates a globally entire grid vector into local patch vectors. Model grid points have its bounded local patch, where the number of local patch vector equals the number of global grid point (the size of local patch size is determined by user option). The shape of each local patch is a cube-like shape vector if the local patch is defined by 3-dimenssional vector space. For instance, when horizontal and vertical grid lengths are l_h and l_v, respectively, the local patch becomes a rectangular cube with $[2l_h + 1] \times [2l_h + 1] \times [2l_v + 1]$ in each grid point. If the coordination of reference model grid point is (i, j, k), 8 grid points to define rectangular cube are $(i \pm l_h, j \pm l_h, k \pm l_v)$. Following description is going to deal with explanations of data assimilation process for each defined local patch because each local patch is independently computed.

In statistical linear estimation theory, the shape of X_a is $L \times N$ matrix composed of an L

dimensional local patch of N ensemble perturbations. The expression of the analyzed state X_a is given by Eq. 2.5 and the updated analysis fields are sequentially projected to the model forecasts in next time step as following Eq. 2.4.

$$X_a = \bar{x}_a + \delta X_a \tag{2.5}$$

where $\bar{x}_a$ is L matrix of analysis ensemble means and δX_a denotes $L \times N$ matrix of the analysis perturbations. Analysis mean, $\bar{x}_a$, is defined by

$$\bar{x}_a = \bar{x}_b + \delta \tilde{x}_a \tag{2.6}$$

where $\bar{x}_b$ is background forecast mean and $\delta \tilde{x}_a$ is analysis increment. The analysis increment determines how the analysis follows observations within a localized patch by comparing observation and background error covariance matrix. The analysis increment is defined by

$$\delta \tilde{x}_a = \delta X_b \tilde{P}_a (\delta Y)^T R^{-1} d \tag{2.7}$$

where δX_b, $\tilde{P}_a$, δY, R, and d are background forecast perturbation, analysis error covariance, forward operated forecast ensemble perturbations, observation error covariance, and observational increment, respectively. The observational increment vector d is difference between observations y_0 and their background ensemble mean counterparts $\overline{H(X_b)}$, where H is possibly a nonlinear observation operator and is replaced to the linearized version of forward operator. A linear observational operator (H) is approximated by a nonlinear observational operator (H) pointed out by Hunt et al. (2007); Miyoshi and Yamane (2007); Szunyogh et al. (2005). The nonlinear observational operator is approximately defined by

$$\mathrm{H}\delta X_b \approx H(\overline{X_b} + \delta X_b) - H(\overline{X_b}) \tag{2.8}$$

to linear forward operator. It is a purpose to project the model grid variable to observation point. Consequently, the forward operated background forecast perturbation is determined by

$$\delta Y = \mathrm{H}(\delta X_b) \tag{2.9}$$

The analysis error covariance is written as

$$\tilde{P}_a = [\delta Y^T R^{-1} \delta Y + (N-1) I/\rho]^{-1} \tag{2.10}$$

where ρ is covariance inflation parameter which enables manual inflation of the analysis error covariance. The parameter helps to avoid the underestimation of the covariance which is common

problem of filter divergence. There are several trials to define this covariance inflation. Regular covariance inflation is the determine the parameter with fixed value through repeated empirical experiments. In this study, we apply a large multiplicative covariance inflation (20% spread inflation) where the inflation parameter is determined to 1.2.

To give weighting according to the distance from the observation, "observation localization" is realized to multiply observational error variance by the inverse of a localization weighting function (based on a Gaussian distribution function). The analysis ensemble perturbations in the $L \times N$ matrix vector, δX_a, are

$$\delta X_a = \delta X_b [(N-1)\tilde{P}_a]^{1/2} \tag{2.11}$$

Therefore, the final analysis ensemble members $\left[x_{a(1)} | \cdots | x_{a(N)}\right]$ in the L dimensional physical space are formed by projection with δX_b:

$$\left[x_{a(1)} | \cdots | x_{a(N)}\right] = \left[\bar{x}_f | \cdots | \bar{x}_f\right] + \delta X_b ([\delta \tilde{x}_a | \cdots | \delta \tilde{x}_a] + \delta X_a) \tag{2.12}$$

When localization patch size (L) is large, the computation efficiency is significantly increased since the multiplication of background forecast perturbation is only taken once in Eq. 2.12. Moreover, the algorithm is more efficient if eigenvalue decomposition is performed for the calculation of analysis error covariance ($\tilde{P}_a$) in Eq. 2.7 (matrix inversion) and Eq. 2.11 (square root filter).

$$\delta Y^T R^{-1} \delta Y + (N-1) I/\rho = U D U^T \tag{2.13}$$

Eqs. 2.7 and 2.11 are written as

$$\tilde{P}_a = U D^{-1} U^T \tag{2.14}$$

$$[(N-1)\tilde{P}_a]^{1/2} = \sqrt{N-1} U D^{-1/2} U^T \tag{2.15}$$

Because the left-hand side matrix of Eq. 2.13 is symmetric, $UU^T = I$ is satisfied. After calculating analysis ensemble states for all independent localized patches, we obtain a representative value at each local patch for the global ensemble analysis. There are several ways to obtain the representative value among components within the localized patch. In this study, the global analysis consists of values at the local patch center within the local patch, which is consistent with previous studies.

Covariance localization is a useful method to moderate spurious error covariance from uncorrelated points with a distance-dependent reduction of error covariance estimates (Hamill, Whitaker, & Snyder, 2001; P. L. Houtekamer & Mitchell, 2001). LETKF method intrinsically contains the localization

through dividing the entire basin grid into independent sub-domain local patches. Even though the LETKF scheme contains the intrinsic localization, Hunt et al. (2007) suggested that the sophisticated localization method with a smooth weighting function in this scheme by weighting the observation error covariance differed from the distance from the local patch center in which the range of weighting function is possibly 0 to 1. The covariance localization via the weighting function within the local patch has a hypothesis in which distant observations have larger errors. It can be realized by multiplying the observation error covariance by the inverse of the smooth weighting function within each local patch. The formulation of the weighting function $w(r_i)$ based on Gaussian function is written as

$$w(r_i) = \exp(-{r_i}^2/2\sigma^2) \tag{2.16}$$

where r_i denotes the distance of ith observation from the local patch center and σ represents a localization scale parameter. In this experiment, we adopt the localization scale parameter value as 30,000meter length scale. The value is an optimized value with several sensitive experiment by determining different localization parameters.

2.5. Evaluation metrics

This study evaluates the assimilated soil moisture estimates according to the assimilation schemes and used satellite retrievals. In this sub-section, we introduce a diagnostic matrix to evaluate the skill of the estimates and the translated innovation. The diagnostics used to analyze the assimilation results are the daily averaged surface and root-zone soil moisture where the vertical depth of them is 0–0.1 m and 0–35 m, respectively.

2.5.1. Anomaly correlation coefficient (ACC)

The assimilation and open loop estimates are assessed against ground based in situ soil moisture measurements. As discussed in Section 2.3, a grid-cell level CDF filtering of soil moisture retrievals mitigates the bias between the rescaled satellite datasets and the JULES model, so that the temporal variation of the observations is primarily assimilated. Hence, the Pearson correlation coefficient is used to evaluate the daily time series of soil moisture estimates on the location of in situ stations, so that one correlation value (R) is yielded per in situ station. In order to focus on the daily variation without low frequency variability, we adopt the anomaly correlations, where the correlation coefficient is calculated with daily anomaly time series removed monthly mean value for each calendar month over the study period.

Based on the Fisher Z transform, we test approximate 95% confidence levels for the anomaly

correlations at in situ sites. These confidence levels depend on the estimated *R* value and on the number of samples (*N*). The approximate 95% confidence intervals are calculated by averaging the 95% confidence intervals and subsequently being divided by the square root of *N*.

2.5.2. Assimilation gain

The concept of assimilation gain is adopted to validate which three components to determine the data assimilation effect on the innovation. In the evaluation of the assimilated soil moisture estimates, the ensemble averaged estimate (Eq. 2.6) is validated with in situ observations. The formulation of the validation is written as

$$Corr(\bar{x}_a, 0) = \frac{1}{T-1}\sum_{i=1}^{T}\left(\frac{\bar{x}_a^i - \overline{\bar{x}_a}}{\sigma_{\bar{x}_a}}\right)\left(\frac{O^i - \bar{O}}{\sigma_O}\right) \tag{2.17}$$

$$Corr(\bar{x}_a, 0) = \frac{\sigma_{\bar{x}_b}}{\sigma_{\bar{x}_a}}\, Corr(\bar{x}_b, 0) + \frac{\sigma_{\delta\tilde{x}_a}}{\sigma_{\bar{x}_a}}\, Corr(\delta\tilde{x}_a, 0) \tag{2.18}$$

$$Corr(\bar{x}_a, 0) - \, Corr(\bar{x}_b, 0) \approx \frac{\sigma_{\delta\tilde{x}_a}}{\sigma_{\bar{x}_a}}\, Corr(\delta\tilde{x}_a, 0) \tag{2.19}$$

where T, O, and σ denotes the days of the validation, in situ observation, and the standard deviation for daily mean $\bar{x}_a$, $\bar{x}_b$, and $\delta\tilde{x}_a$, respectively. The skill improvement of surface and root-zone soil moisture is computed by the difference of the temporal correlation coefficient between the assimilated soil moisture estimates and the open loop model. The skill from the open loop is represented by $Corr(\bar{x}_b, 0)$ in Eq. 2.18 and the daily variability of $\sigma_{\bar{x}_b}$ and $\sigma_{\bar{x}_a}$ is almost comparable. Therefore, Eq. 2.18 is approximately represented by Eq. 2.19. In the Eq. 2.7, $\delta\tilde{x}_a$ (analysis increment) is separated by $\delta X_b \tilde{P}_a (\delta Y)^T R^{-1}$ (Kalman gain) and d (observational increment), where the analysis increment is the inner product of the Kalman gain and the observational increment. The inner product is calculated with respect to the dimension of the number of assimilated observations. Hence, the analysis increment term is written as

$$\delta\tilde{x}_a \approx K \times N_{sat} \times d \tag{2.20}$$

where K and N_{sat} denotes the Kalman gain and the number of assimilated observations, respectively. In the data assimilation process, the Kalman gain and the number of observations is almost constant. Therefore, we can rewrite the right term of Eq. 2.19 as

$$\frac{\sigma_{\delta\tilde{x}_a}}{\sigma_{\bar{x}_a}}\, Corr(\delta\tilde{x}_a, 0) = \frac{K \times N_{sat} \times \sigma_d}{\sigma_{\bar{x}_a}} Corr(d, 0) \tag{2.21}$$

In other words, the skill improvement of the soil moisture through the data assimilation is determined by (1) the quality of remote sensing retrievals, (2) the Kalman gain appropriating weight to the observations, and (3) the number of samples in the assimilation process. In the Eq. 2.21, σ_d is almost similar with each experiment because the original satellite observation is processed in the bias correction on the open loop model, but $\sigma_{\bar{x}_a}$ is different for each assimilation experiment. Therefore, in this study, we develop a diagnostic matrix to understand the skill improvement through the data assimilation with LETKF scheme. It is referred to "Assimilation Gain" which consists of the skill difference between the satellite and the open loop model ΔR, the Kalman gain, and the number of assimilated observational sample N_{sat}, and they are proportional to the skill improvement in the assimilation. The assimilation gain is written as

$$Assimilation\ Gain = \Delta R_{sat} \times K \times N_{sat} \tag{2.22}$$

$$\Delta R_{sat} = R_{sat} - R_{openloop} \tag{2.23}$$

$$K = \frac{\sum_{t=1}^{N_{days}} \left[\frac{E_b(t)}{E_b(t)+E_o(t)} \right]}{N_{days}} \tag{2.24}$$

$$N_{sat} = \frac{\sum_{t=1}^{N_{days}} \sum_{i=1}^{n} w(r_i)_t}{N_{days}} \tag{2.25}$$

where R_{sat} and $R_{openloop}$ are the temporal anomaly correlation of remotely sensed retrievals (y_0) and the open loop ($\overline{H(X_b)}$), respectively (Eq. 2.23). For the Kalman gain, $E_b(t)$ and $E_o(t)$ is error covariance of model background and the observation of daily mean soil moisture, respectively, and N_{days} denotes the number of day over the research period. High values of the Kalman gain implies that the analysis soil moisture is closer to the observation than to the background and the value is bounded from 0 to 1. The number of assimilated observational sample is defined as the daily mean of the summing weighting function (Eq. 2.24) within the local patch for the study period (Eq. 2.25). The assimilation gain is only used to investigate the sensitivity of the data assimilation according to different retrievals with the LETKF scheme. When comparing the impact of assimilation schemes, the Kalman gain is enough to evaluate their assimilation sensitivities if a homogeneous observation is used.

2.5.3. Drought index

To evaluate the impact of assimilated soil moisture on the drought monitoring, this study adopts a drought index which consists of representative meteorology variables and soil moisture state. For the meteorology variables, precipitation and surface temperature is used in this drought index and they are

normalized from 0 to 1, scaled by their maximum and minimum values at each pixel, which considers the potential maximum and minimum from long-term analysis. For each component, we use precipitation condition index (PCI), temperature condition index (TCI), and soil moisture condition index (SMCI) and they are written as

$$PCI = \frac{Pr - Pr_{min}}{Pr_{max} - Pr_{min}} \tag{2.26}$$

$$TCI = \frac{Temp_{max} - Temp}{Temp_{max} - Temp_{min}} \tag{2.27}$$

$$SMCI = \frac{SoilM - SoilM_{min}}{SoilM_{max} - SoilM_{min}} \tag{2.28}$$

where Pr, $Temp$, and $SoilM$ represent a few daily or monthly average of precipitation and surface temperature and surface soil moisture, respectively, and their subscribed values are the minimum and maximum value from the long-term analysis corresponding the average period. The closer the drought index is to zero, the more severe the drought phenomenon.

2.5.4. Soil moisture-temperature coupling index

The process of land-atmospheric interaction is complex and variable. During drought or heatwave events, the interaction is strong with soil moisture deficit and anomaly high sensible heat flux at land surface. Understanding the energy balance by the assimilated soil moisture estimates at land surface is important for representing realistic land-atmospheric interaction in the data assimilation. To diagnose the land-atmospheric interaction, this study adopts a soil moisture-temperature coupling index (D. Miralles, Van Den Berg, Teuling, & De Jeu, 2012). The coupling index (Π) is the estimation of two energy balances of $H = R_n - \lambda E$ and $H_p = R_n - \lambda E$ (actual evaporation (E) and potential evaporation (E_p) compared to net radiation(R_n)) and their difference in the variability of near-surface temperature (T), and its short time-scales variability is defined as

$$\Pi_t = \frac{T_t - \bar{T}}{\sigma_T}\left(\frac{H_t - \bar{H}}{\sigma_H} - \frac{H_{p,t} - \overline{H_p}}{\sigma_{H_p}}\right) \tag{2.29}$$

where $\bar{T}$, $\bar{H}$ and $\overline{H_p}$ are the means of the long-term series of T, H and H_p, and σ_T, σ_H and σ_{H_p} are the standard deviations of them, correspondingly. Their subscribed value (t) denotes everyday time. This coupling metrics is available when all components have their long-term data. The variables from the data assimilation experiment do not have long-term analysis, so that we adapt the land surface conditions simulated by the long-term open loop model in the coupling analysis of the data assimilation

experiment. It is not a problem to calculate using this method, because raw satellite datasets are corrected against the open loop climatology. It means that the climatology from the assimilation experiment and the open loop model is the same regardless of whether the satellite datasets are assimilated, but their variability is improved. However, in this study, we cannot calculate the observed soil moisture-temperature coupling index because there is no long-term in situ observation enough to be analyzed.

2.6. Experiment designs

As addressed in Reichle, Crow, and Keppenne (2008), the meteorological boundary forcing variables and the model prognostic variable are perturbed using the perturbations parameter values to explain uncertainties related to errors in LSM physics and atmospheric boundary forcing variables. These uncertainties are contained in the model error w_b referred in Eq. 2.4. For explaining uncertainties of the model physics and the boundary forcing inputs, we perturb random numbers to the radiation and rainfall forcing variables, and the JULES model prognostic variable related to soil moisture. As displayed in Table 2.4, normally distributed additive perturbation is applied to the soil moisture prognostic variable and to the longwave radiation forcing and lognormally distributed multiplicative perturbation is applied to the precipitation and the shortwave radiation forcing, where the ensemble mean of additive and multiplicative perturbation is 0 and 1, respectively. A first-order auto-regressive model (AR1) time series correlation scales are applied for all perturbation fields. In addition, cross correlations ensure that the meteorological variables represent a realistic balance between the perturbed forcing variables. Under this assumption, a positively perturbed downward shortwave radiation tends to be related with negative perturbations of the downward longwave radiation and the precipitation, and vice versa. A detailed description of perturbing the meteorological boundary forcing variables and the model prognostic variable is provided in Reichle et al. (2008).

There are several parameters to determine the soil moisture data assimilation system. As discussed in Eq. 2.19, the localization scale parameter is used to weight observations within a localized patch depending on the distance between the location of the observations and local patch center. Moreover, the LETKF method can control analysis perturbation through increasing the inflation parameter in the analysis error covariance (Eq. 2.10). To optimize the land data assimilation system based on the LETKF scheme, we perform several sensitivity experiments according to localization scale parameter and covariance inflation parameter. The specific description for experiment design to examine these parameters is given in Table 2.5 and Table 2.6 in which factors modified in sensitivity experiments are highlighted.

This study mainly focuses on the two aspects: (1) the impact of two different data assimilation schemes (EnKF and LETKF) on soil moisture estimates and (2) the sensitivity of active and passive microwave satellites in the soil moisture data assimilation. In order to investigate those aspects this study performs two different data assimilation experiments. The specific experiment information for each experiment is described in Table 2.7 and Table 2.8 where the main points in each experiment are marked with red color.

Table 2.4 Parameters for perturbations to near surface atmospheric boundary forcing variables and JULES model prognostic value associated with the soil moisture

Perturbation variables	Additive (A) / Multiplicative (M)	Standard deviation	AR1 Time series correlation scale
Precipitation	M	0.5	1 day
Downward shortwave (SW)	M	0.3	1 day
Downward longwave (LW)	A	50 W m^{-2}	1 day
Moisture concentration	A	0.002 m^3 m^{-3}	3 hours

Table 2.5 Experiment design for evaluating the impact of localization scale parameter on the soil moisture data assimilation

	Information	Reference
Land Surface Model	JULES	Best et al. (2011)
Atmospheric boundary forcing	JRA-55 reanalysis corrected by 6-hourly GSMaP rainfall and monthly CMAP precipitation datasets	Aonashi et al. (2009); Kobayashi et al. (2015); Kubota et al. (2007); Ushio et al. (2003); Ushio et al. (2009); Xie and Arkin (1997)
Soil moisture observation	Level-2 ASCAT C-band active (radar) Level-2 SMAP L-band passive (radiometer)	Chan et al. (2016); (Wagner et al., 2013)
Data assimilation scheme	Local Ensemble Transform Kalman Filter (LETKF)	Evensen (2003); Hunt et al. (2007); Miyoshi and Yamane (2007)
Resolution	0.5°×0.5° (~50km) 4 layers (0.1, 0.25, 0.65, and 2 meters) 3-hourly DA cycle	
Research period	May-September, 2016	
Localization patch size	sig=30km / 100km / 190 km	
Inflation parameter	ρ=1.2	
Ensemble sizes	12	

Table 2.6 Experiment design for evaluating the impact of covariance inflation parameter on the soil moisture data assimilation

	Information	**Reference**
Land Surface Model	JULES	Best et al. (2011)
Atmospheric boundary forcing	JRA-55 reanalysis corrected by 6-hourly GSMaP rainfall and monthly CMAP precipitation datasets	Aonashi et al. (2009); Kobayashi et al. (2015); Kubota et al. (2007); Ushio et al. (2003); Ushio et al. (2009); Xie and Arkin (1997)
Soil moisture observation	Level-2 ASCAT C-band active (radar) Level-2 SMAP L-band passive (radiometer)	Chan et al. (2016); (Wagner et al., 2013)
Data assimilation scheme	LETKF	Evensen (2003); Hunt et al. (2007); Miyoshi and Yamane (2007)
Resolution	0.5°×0.5° (~50km) 4 layers (0.1, 0.25, 0.65, and 2 meters) 3-hourly DA cycle	
Research period	May-September, 2016	
Localization patch size	sig=30km	
Inflation parameter	ρ=1.0 / 1.2 / 1.5 / 1.7	
Ensemble sizes	12	

Table 2.7 Experiment design for evaluating the impact of two different assimilation schemes (EnKF vs. LETKF) on the soil moisture data assimilation

	Information	Reference
Land Surface Model	JULES	Best et al. (2011)
Atmospheric boundary forcing	JRA-55 reanalysis corrected by 6-hourly GSMaP rainfall and monthly CMAP precipitation datasets	Aonashi et al. (2009); Kobayashi et al. (2015); Kubota et al. (2007); Ushio et al. (2003); Ushio et al. (2009); Xie and Arkin (1997)
Soil moisture observation	Level-2 SMAP L-band passive (radiometer)	Chan et al. (2016)
Data assimilation scheme	EnKF / LETKF	Evensen (2003); Hunt et al. (2007); Miyoshi and Yamane (2007)
Resolution	0.5°×0.5° (~50km) 4 layers (0.1, 0.25, 0.65, and 2 meters) 3-hourly DA cycle	
Research period	May-September, 2016	
Localization patch size	3°×3° (150km), sig=30km	
Inflation parameter	ρ=1.2	
Ensemble sizes	12	

Table 2.8 Experiment design for evaluating the impact of remote sensing soil moisture retrievals (ASCAT vs. SMAP) on the soil moisture data assimilation

	Information	Reference
Land Surface Model	JULES	Best et al. (2011)
Atmospheric boundary forcing	JRA-55 reanalysis corrected by 6-hourly GSMaP rainfall and monthly CMAP precipitation datasets	Aonashi et al. (2009); Kobayashi et al. (2015); Kubota et al. (2007); Ushio et al. (2003); Ushio et al. (2009); Xie and Arkin (1997)
Soil moisture observation	Level-2 ASCAT C-band active (radar) Level-2 SMAP L-band passive (radiometer) Level-2 SMAP + ASCAT joint dataset	Chan et al. (2016); Wagner et al. (2013)
Data assimilation scheme	LETKF	Evensen (2003); Hunt et al. (2007); Miyoshi and Yamane (2007)
Resolution	0.5°×0.5° (~50km) 4 layers (0.1, 0.25, 0.65, and 2 meters) 3-hourly DA cycle	
Research period	May-September, 2016	
Localization patch size	3°×3° (150km), sig=30km	
Inflation parameter	ρ=1.2	
Ensemble sizes	12	

Chapter 3

3. Implication of land data assimilation system

3.1. Background

Soil moisture plays an important role over the partition of global land surface water and energy. To derive global soil moisture distributions, there are two alternative data sources: (1) the offline simulation of LSM forced with observed precipitation, radiation, and other atmospheric variables and (2) remotely sensed near-surface soil moisture. However, the land surface dataset driven by the LSM offline simulation is subject to several limitations such as the model physical process, the specification of vegetation and soil parameters, and the forcing data. Moreover, satellite retrievals have a limitation for their temporal and spatial coverages and represent only near-surface soil moisture even though root zone soil moisture is also important.

The assimilation of passive and active microwave soil moisture data merges satellite retrievals and model soil moisture and provides optical global soil moisture estimates. There are some previous studies to perform the soil moisture assimilation in which all of studies commonly adopts the EnKF scheme (Draper et al., 2012; Kolassa et al., 2017; Lievens et al., 2017; Q. Liu et al., 2011), although the scheme is not feasible in large modeling domains or high spatial resolution of the land data assimilation and a lot of observations due to the limitation of computation resources. However, in recent days, the available remotely sensed soil moisture datasets are increased, so that this study develops the framework of land data assimilation system with the LETKF scheme.

Based on the data assimilation system, this study examines the dependency of active and passive remote sensing soil moisture retrievals and compares computing efficiency to the results of the EnKF. Q. Liu et al. (2011) demonstrated that assimilating AMSR-E retrievals increases soil moisture skills by $\Delta R \approx 0.08$ compared to the open loop of Catchment LSM. Assimilating ASCAT satellite data improves soil moisture skills by $\Delta R \approx 0.1$, which is comparable to the results of AMSR-E (Draper et al., 2012). The results of jointly assimilating ASCAT and AMSR-2 observations is not much different with the individually assimilated experiments by both of retrievals, which means that radar and radiometer satellites are not much complementary. SMAP satellite has large antenna (6m diameter) and the operating active and passive microwave channels (centered at 1.26 GHz and 1.41 GHz, respectively). The characteristics help a higher soil moisture measurement sensitivity compared to previous sensors.

The combined active (radar) and passive (radiometer) sensors tends to better control radio frequency interference (RFI) (Piepmeier et al., 2014) and achieve measurement accuracy and spatial resolution. However, the active sensor does not possible due to sudden failure of the radar. Pan et al. (2016) suggested that the SMAP provides significant added value and positive potential for data assimilation which is better that the result of VIC (Variable Infiltration Capacity) comparted to the ground based in situ observations. Actually, Lievens et al. (2017) demonstrated the improvement of soil moisture skills by $\Delta R \approx 0.06$ by the assimilation of the SMAP data.

Beside these previous studies related with the soil moisture assimilation, this study has a purpose to evaluate the impact of assimilation schemes and remote sensing retrievals on the soil moisture estimates. In this section, we will compare the results from two different data assimilation scheme between LETKF (in this study only) and EnKF (in most previous studies) and evaluate the results of assimilated soil moisture skill with ASCAT, SMAP, and interactive ASCAT & SMAP, respectively.

3.2. Optimization of LETKF scheme

3.2.1. Localization scale parameter

As described in Table 2.5, we conduct several sensitivity experiments according to the localization scale parameter with LETKF based assimilation system. The experiments are repeatedly performed with ASCAT and SMAP remotely sensed satellite datasets. This study evaluates the soil moisture skill assimilated with different localization scale parameter setting compared to ground based in situ soil moisture measurements.

Figure 3.1 represents the mean skill of surface and root-zone soil moisture from the assimilation experiments and the open loop model. The skill is categorized by MODIS land cover and its 95% confidence level is tested by the Fisher Z transform method. The skill of soil moisture simulated by the open loop is approximately 0.5 and the results from the assimilation experiments tends to show better skill score even though the improvement depends on the localization scale parameter. The skill increasing (up to ~0.15) is mostly significant on grassland cover and the improvement is not much dominant over mixed land cover and cropland. It is commonly shown in the results of ASCAT and SMAP and the overall skill improvement of SMAP is larger than those in ASCAT. As increasing the localization parameter, the soil moisture skill difference between the assimilation and the open loop experiment tends to be decreased because uncorrelated observations within each local patch are influenced on the innovation. It is commonly revealed in the result of ASCAT and SMAP (Fig. 3.2).

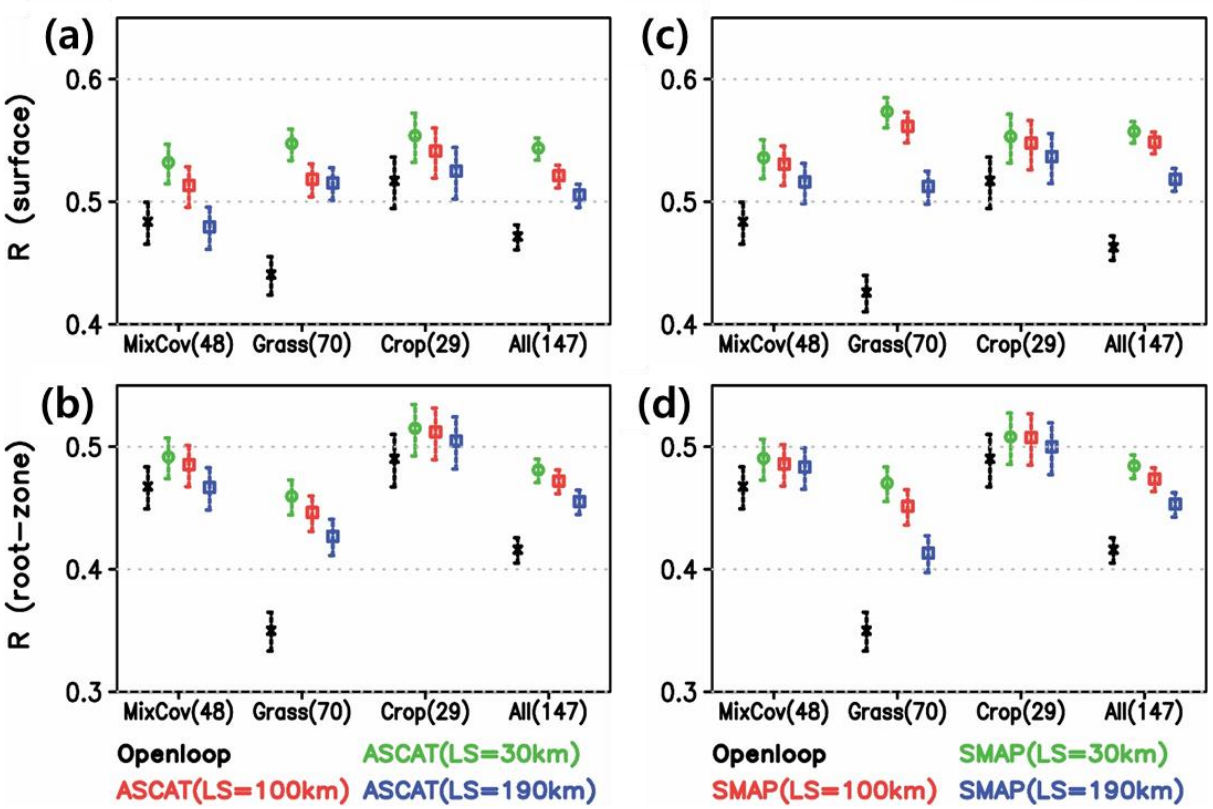

Figure 3.1 (a, c) Surface and (b, d) root-zone soil moisture skill from the soil moisture assimilation and the open loop run, averaged across categorized MODIS land cover, where errors represent 95% confidence intervals. The results of ASCAT and SMAP are on left and right column, respectively. The number denoted on each land cover represents in situ sites used to evaluate the skill.

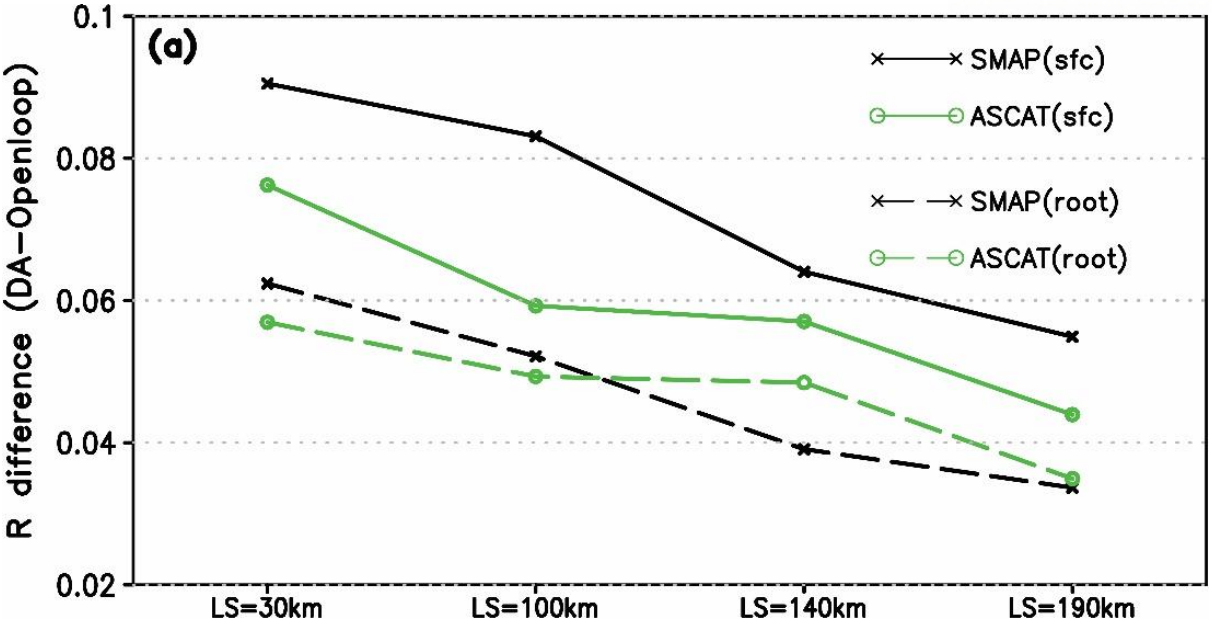

Figure 3.2 Mean skill difference between the data assimilation and the open loop simulation over all in situ sites according to different localization scale parameters. Black and green lines represent the results of SMAP and ASCAT, respectively, and solid and dashed lines display the results of surface and root-zone soil moisture.

3.2.2. Covariance inflation parameter

As described in Table 2.6, we conduct several sensitivity experiments with different covariance inflation parameters with LETKF based data assimilation system. The experiments are respectively performed with ASCAT and SMAP remotely sensed satellite datasets. This study evaluates the soil moisture skill assimilated with different inflation parameters validated with in situ soil moisture measurements.

Figure 3.3 shows the averaged skill of surface and root-zone soil moisture from the assimilation experiments and the open loop model across categorized MODIS land cover. All of soil moisture assimilation experiments represent the significant skill improvement, but there is large sensitivity of the soil moisture skill depending on the covariance inflation parameter. For instance, in the experiments with ASCAT retrieval, the soil moisture skills are decreased as increasing the perturbation of background states, but the result from the SMAP assimilation experiments shows the opposite by the inflation parameter. It means that the impact of controlling inflation parameter on the skill of assimilated soil moisture estimates shows the dependency of each satellite. Moreover, the sensitivity is also shown by the land cover classes. In general, the improvement of soil moisture skill is reduced when the inflation parameter is increased in the assimilation experiments using ASCAT observation, and the result from the assimilation experiments using SMAP observation is shown oppositely (Fig. 3.4).

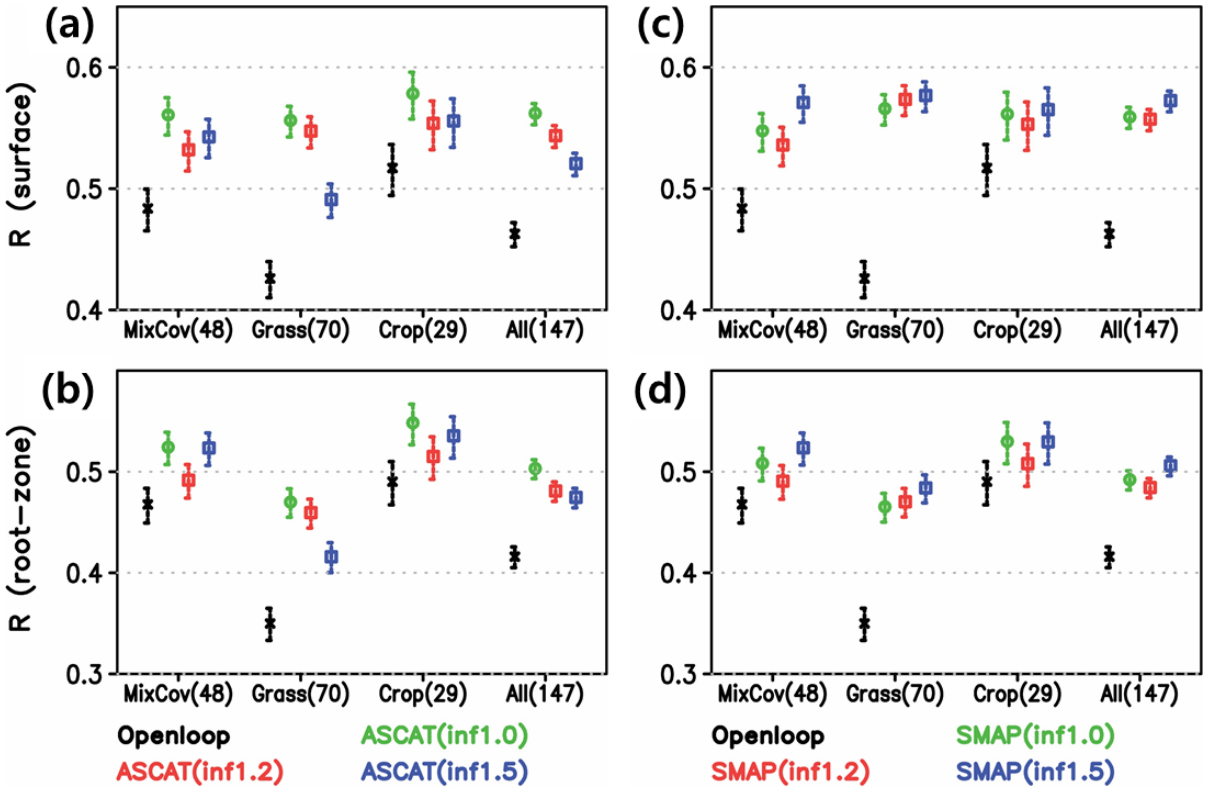

Figure 3.3 (a, c) Surface and (b, d) root-zone soil moisture skill from the soil moisture assimilation with different inflation parameter of 1.0 (green), 1.2 (red), and 1.5 (blue) and the open loop run, averaged across categorized MODIS land cover, where errors represent 95% confidence intervals. The results of ASCAT and SMAP are on left and right column, respectively.

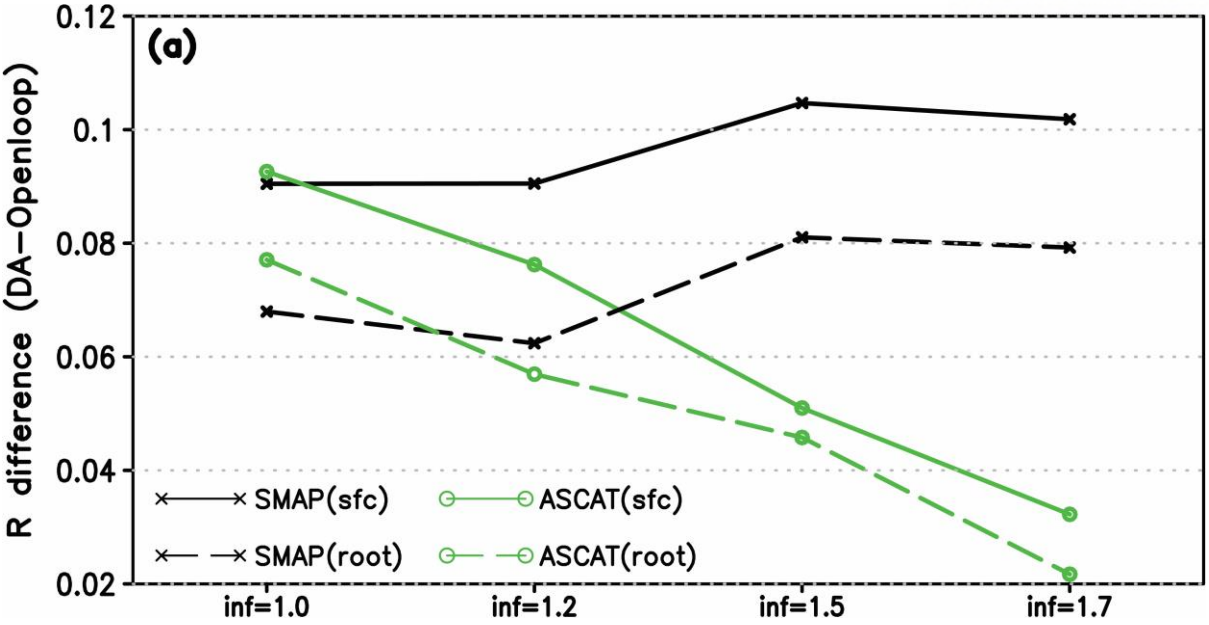

Figure 3.4 Mean skill difference between ASCAT and SMAP single-sensor assimilation experiments and the open loop simulation over all in situ sites according to different covariance inflation parameters. Black and green lines represent the results of SMAP and ASCAT, respectively, and solid and dashed lines display the results of surface and root-zone soil moisture.

3.3. Impact of assimilation schemes on the data assimilation system

In this sub-section, we evaluate the impact of major features of LETKF on the computational efficiency and the skill of assimilated soil moisture estimates. The specific description for the experiments is given in Table 2.7.

3.3.1. Computational efficiency

The main difference of the LETKF scheme compared to the EnKF is the decomposition of global analysis domain into a number of independent sub-domains, which is referred to the localization. Hence, the EnKF should calculate large size of matrix with three dimensions of model variables, a number of observations, and ensemble members over the entire domain at once. In contrast, the LETKF is operated within decomposed independent sub-domains and repeats the assimilation process in the sequential localized patch. For instance, when we conduct the data assimilation over the North America (5733 model variables) with 12-member forecast ensembles and 339-number observations, CPU computation time is consumed 38 (LETKF) and 115 (EnKF) seconds (Fig. 3.5), respectively. 3 times larger computing resources are demanded in the EnKF compared to the LETKF. When the larger number of observations about 11268 are assimilated, CPU computation time needs 83 (LETKF) and 15569 (EnKF) seconds, respectively, where approximately 180 times more computing resources are demanded in the non-localized assimilation scheme. In other words, the LETKF is much efficient to apply the data assimilation even if the assimilation skill is not much different compared to the skill of the EnKF.

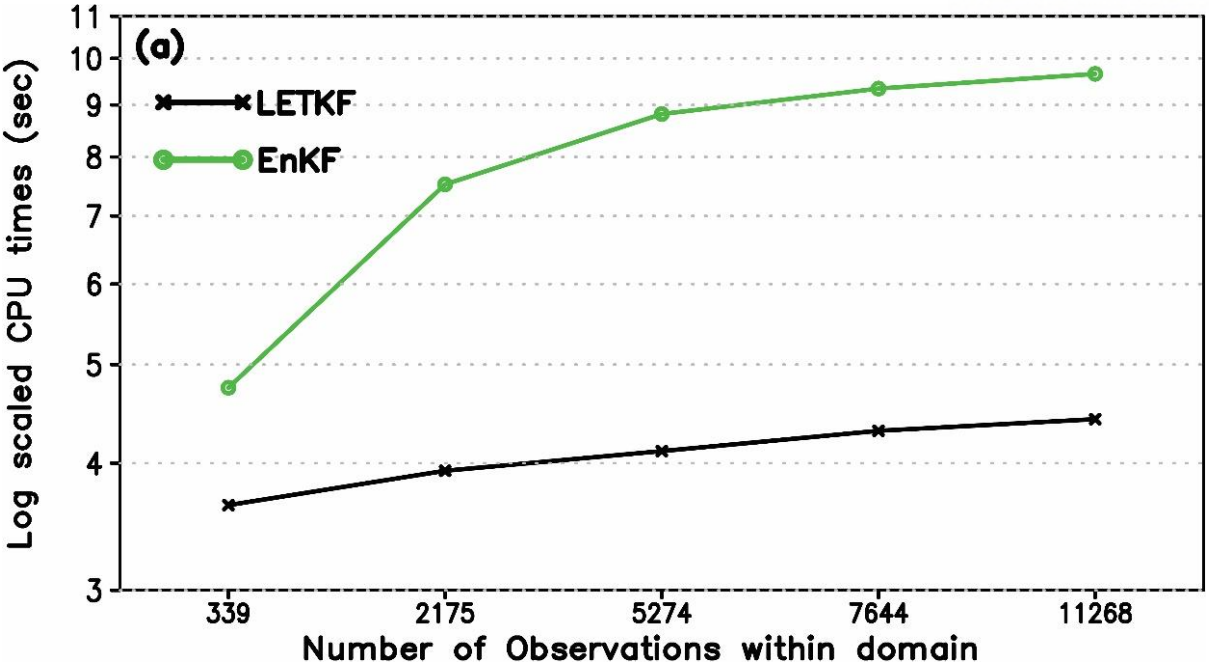

Figure 3.5 CPU times to calculate the analysis fields with a different number of observations within the North America domain (130°W–65°W, 25°N–55°N) using the EnKF and the LETKF schemes.

3.3.2. Evaluation of soil moisture skill

As we discussed the characteristics of LETKF and EnKF in Section 2.4, another main difference between both schemes is the transform which enables to covert the background and observation error covariance to the analysis perturbation through the square root filtering method. The analysis error covariance matrix included in the LETKF scheme has the covariance inflation parameter which can control the analysis perturbation. The underestimating background ensemble spread is a fatal flaw of the original formulation of EnKF, but the inflation parameter helps to moderate the chronic problem of non-variational data assimilation. Its impact on the soil moisture skill is also described in Section 3.2.2. In this section, we will compare the soil moisture skill assimilated by EnKF and LETKF methods only using SMAP retrieval.

Figure 3.6 represents the skills of surface and root-zone soil moisture estimates from the SMAP soil moisture assimilation experiments and the open loop and they are also categorized by MODIS land cover. The assimilation experiments with SMAP data are conducted with different assimilation schemes such as EnKF and LETKF methods and we perform two sets of LETKF data assimilation experiments according to the inflation parameter of 1.0 and 1.5. Overall soil moisture skills are comparable between assimilation experiments even though the assimilation schemes used in these experiments are entirely different. The improvement through the assimilation is mostly significant over the grassland cover in which the soil moisture skill simulated by the open loop is worst. However, there are slight skill improvement of soil moisture estimates through emphasizing background error covariance by increasing the inflation parameter and the soil moisture skill applied by background error inflation tends to be better representation compared to the result of EnKF. The soil moisture skill assimilated by LETKF scheme with increased inflation parameter is higher than the results from the EnKF assimilation, which is accompanied by the Kalman gain change (Fig. 3.7). The skill improvement has a land cover dependency, where the innovation is mostly dominant in mixed land cover regions and related to the increased background perturbation.

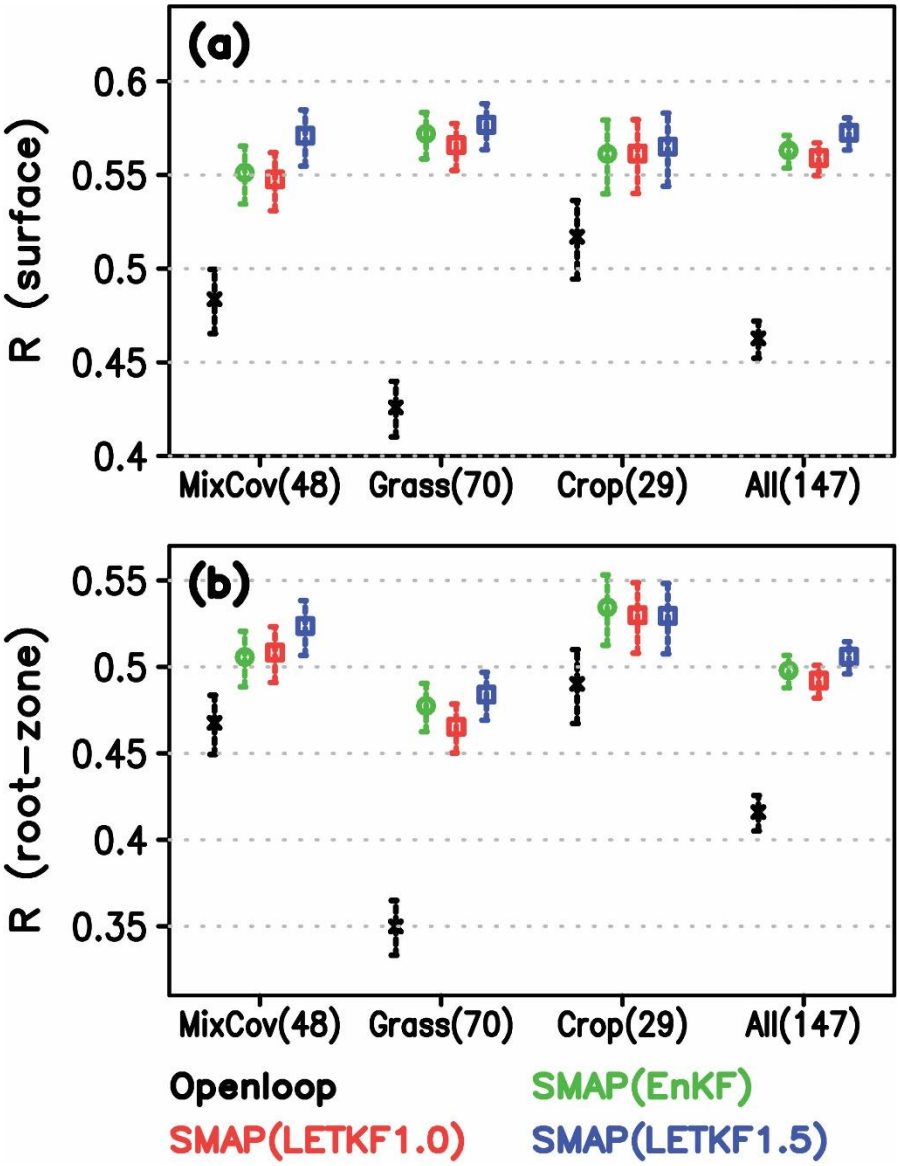

Figure 3.6 (a) Surface and (b) root-zone soil moisture skill from the SMAP soil moisture assimilation experiments and the open loop run, averaged across categorized MODIS land cover, where errors represent 95% confidence intervals. The SMAP assimilation experiments are performed with different assimilation schemes: EnKF (green), and LETKF with 1.0 (red) and 1.5 (blue) covariance inflation parameter.

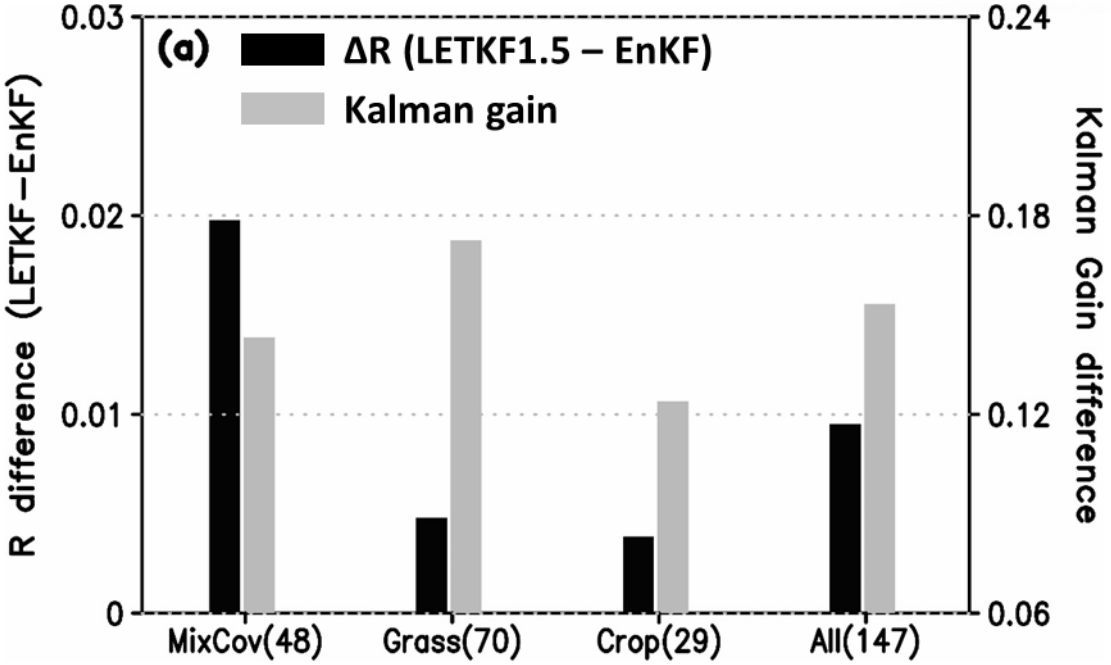

Figure 3.7 Mean values of assimilated surface soil moisture skill (black bar) and Kalman gain (gray bar) difference between SMAP assimilation experiments with EnKF and LETKF scheme.

3.4. Impact of assimilated remote sensing retrievals on soil moisture skill

As described in Table 2.8, we conduct three data assimilation experiments of different remote sensing soil moisture retrievals with the LETKF scheme. In these experiments, level-2 ASCAT and SMAP soil moisture datasets are used in the assimilation and the joint assimilation with both observations is also performed. Based on the optimization of the soil moisture data assimilation system described in Section 3.2, we adopt the best configuration of the LETKF scheme. The skill of the assimilated soil moisture estimates is evaluated with the anomaly correlation coefficient as introduced in Section 2.5.1.

3.4.1. Evaluation of soil moisture estimates assimilated with remote sensing retrievals

To evaluate the impact of assimilated remote sensing retrievals on soil moisture skill, this study conducts three set of soil moisture assimilation experiments with several combinations of ASCAT and SMAP satellites as described in Table 2.8. The correlation coefficient of daily evolution of soil moisture estimates from three data assimilation experiments and the open loop, compared to ground based in situ observations, represents their quality of assimilated soil moisture states.

Before we evaluate the performance of the soil moisture assimilation experiments, the quality of remotely sensed soil moisture retrievals (ASCAT and SMAP) is validated by the ground-based measurements, where generally SMAP satellite tends to reveal the higher skill score than those in ASCAT satellite (Fig. 3.8a). Especially, the skill difference between ASCAT and SMAP retrievals is the largest of $\Delta R=0.17$ in the cropland class. In addition, when we compare the soil moisture skill from both satellites to the open loop simulation, the skill differences between remote sensing retrievals and the open loop model shows that SMAP is large ($\Delta R=0.15$), and ASCAT is not much ($\Delta R=0.02$). The open loop model shows better soil moisture skill than the result of ASCAT in the vegetated mixed land cover and the cropland classes, but SMAP represents better skill in all land cover classes than those in the open loop. The spatial distribution of soil moisture skill difference between assimilated ASCAT and SMAP satellite and the open loop simulation over the North America is also shown in Fig. 3.9. The quality of ASCAT satellite tends to show better than the open loop model over Utah, Colorado, and Alabama states, but the satellite quality is not much good in other states. The skill of soil moisture simulated by the open loop model is strongly tided by meteorological variables. In particular, the precipitation mainly contributes to determine the accuracy of the soil moisture which is well explained by the mean skill of precipitation (Fig. 3.8b).

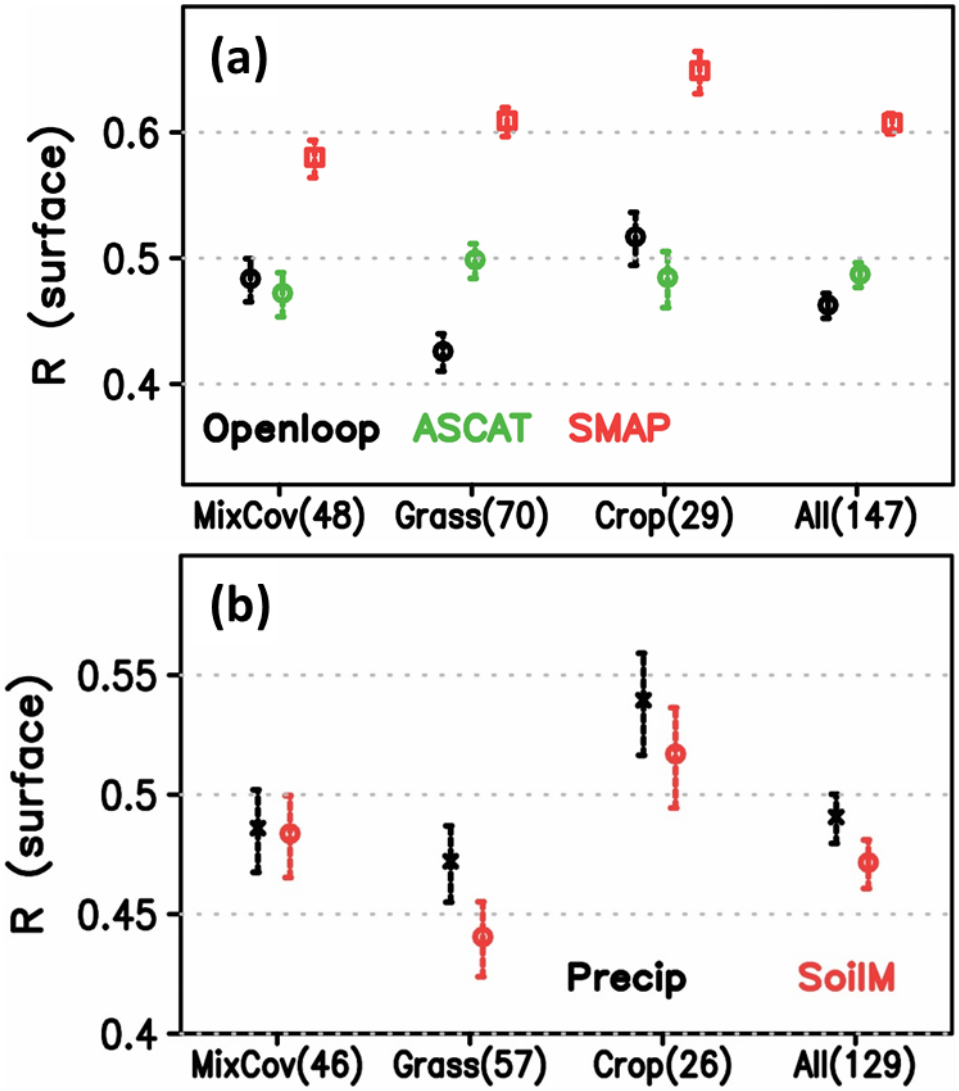

Figure 3.8 (a) Mean skill of ASCAT (green) and SMAP (red) soil moisture retrievals and the open loop model (black) averaged across the vegetated mixed land cover, the grassland and the cropland classes, where errors represent 95% confidence intervals. (b) Soil moisture skill from the open loop model (black) and precipitation (red) from atmospheric boundary forcing. The skills are evaluated with in situ measurements over the North America and the comparison between soil moisture and precipitation skill is done only where both observations exist.

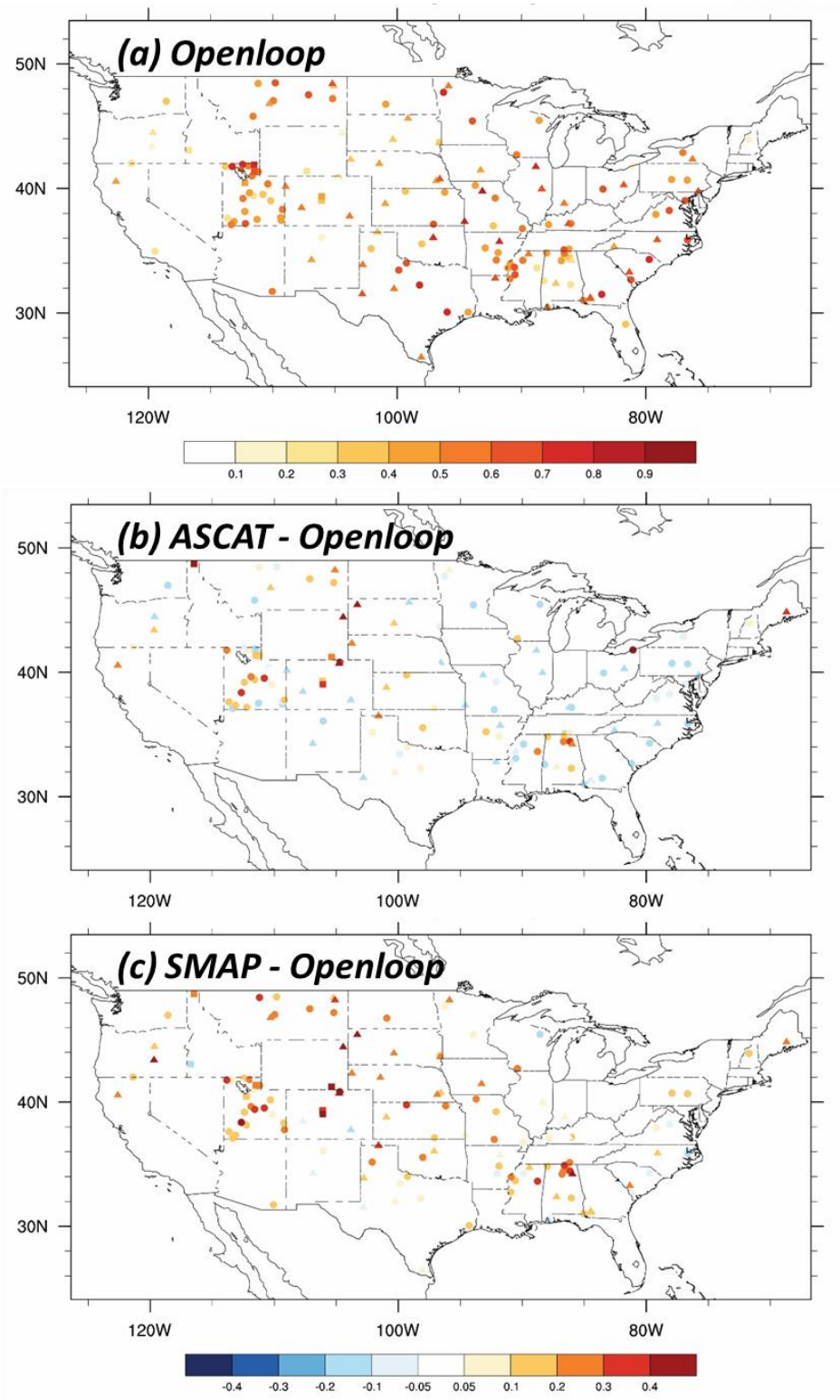

Figure 3.9 The spatial distribution of (a) soil moisture skill from the open loop model and the skill difference between (b) ASCAT and (c) SMAP satellite and the open loop simulation, evaluated by ground based in situ stations over the North America. In these difference maps, red color indicates a higher skill of the satellite observation and blue color indicates a higher skill of the open loop model.

Figure 3.10 represents mean skill of surface and root-zone soil moisture estimates from the single-sensor assimilation experiments (of ASCAT or SMAP) and the combined assimilation experiment (ASCAT and SMAP) with the optimized LETKF scheme and the open loop run, averaged across categorized MODIS land cover. Prior to the assimilation, the soil moisture skill simulated by the open loop has significantly lower over the grassland cover than over the vegetated mixed cover or cropland categories. The single-sensor assimilation experiments yield very similar improvements (ΔR of ASCAT and SMAP assimilation experiments is 0.08 and 0.1, respectively) to the mean skill, although the SMAP assimilation experiment tends to show the better representation of surface soil moisture over the vegetated mixed cover class. However, the results of the two single-sensor retrieval products is not much different since the skill of the open loop is already saturated over the cropland class. Rather, the skill difference of surface and root-zone soil moisture from the two experiments is more significant in the grassland class where the performance of the open loop is worst among the other land cover classes. Nevertheless, the joint assimilation significantly exceeds the mean correlation from the single-sensor assimilation experiments by ΔR=0.12. In other words, the joint assimilation of ASCAT and SMAP satellites performs best, demonstrating the complementary value of active and passive remoted sensed datasets. The value of skill improvement via the data assimilation is comparable to several previous studies (Draper et al., 2012; Kolassa et al., 2017; Lievens et al., 2017) where the skill improvement of surface and root-zone soil moisture through the assimilating active and passive satellites is approximately up to ~0.1.

In order to investigate the skill improvement, we adopt the concept of assimilation gain which is introduced in Section 2.5.2. The assimilation gain consists of the skill difference between the satellite and the open loop, the Kalman gain, and the number of assimilated observational sample. Figure 3.11 represents mean values of soil moisture skill difference and components of the assimilation gain from three assimilation experiments over the North America and their spatial distributions are shown in Figure 3. 12. The assimilation gain reliably explains the skill improvement in which they are proportional with each other. This result is shown by the combination of the three components constituting the assimilation gain. For instance, in the single-sensor assimilation experiments, the higher assimilation gain value in the SMAP experiment can be largely explained by the relative skill of the satellite rather than the number of assimilated observations and the Kalman gain. In the comparison of the combined assimilation experiment to the SMAP assimilation, the higher assimilation gain value in the combined experiment can be explained by the number of assimilated observations rather than the relative skill of the satellite and the Kalman gain. The results are also represented by radar plot with three components in which the triangle area is proportional to the assimilation gain (Fig. 3.13). Based on this analysis, the assimilation gain can be used as a parameter that contributes to the potential predictability improvement.

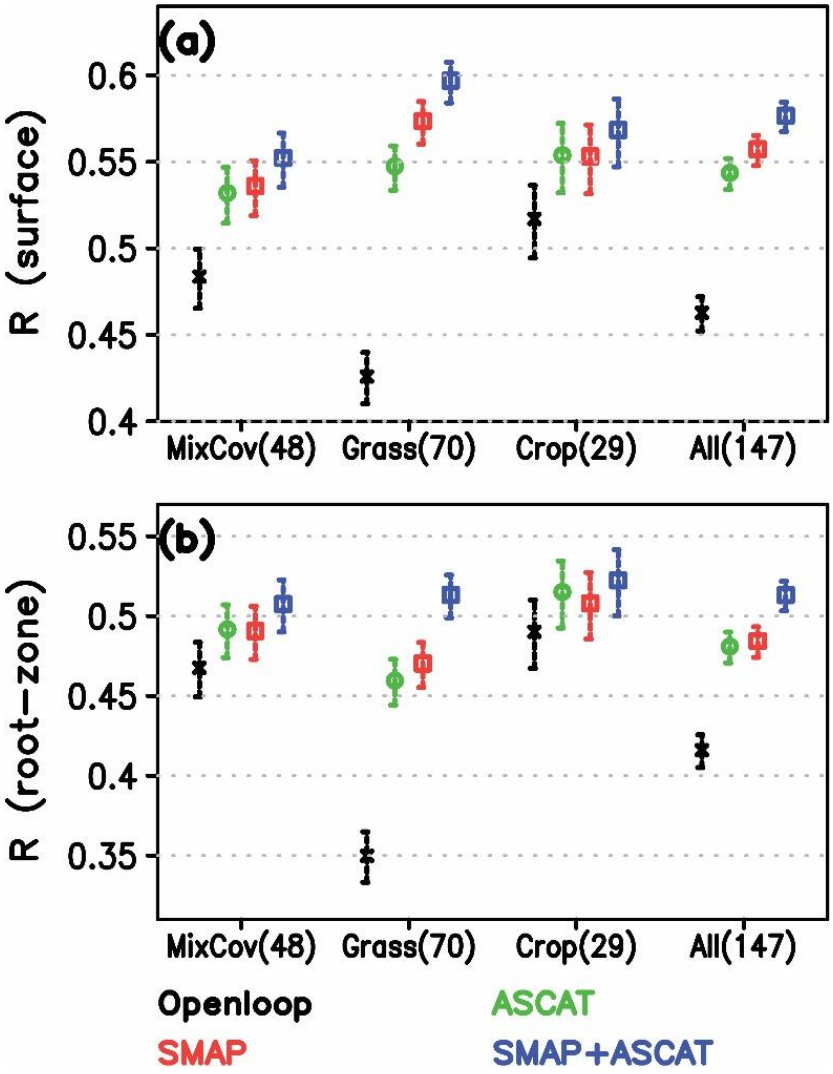

Figure 3.10 (a) Surface and (b) root-zone soil moisture skill from ASCAT (green), SMAP (red), and joint retrieval (blue) assimilation experiments with LETKF scheme and the open loop run, averaged across categorized MODIS land cover, where errors represent 95% confidence intervals. The soil moisture skills are evaluated with in situ measurements over the North America.

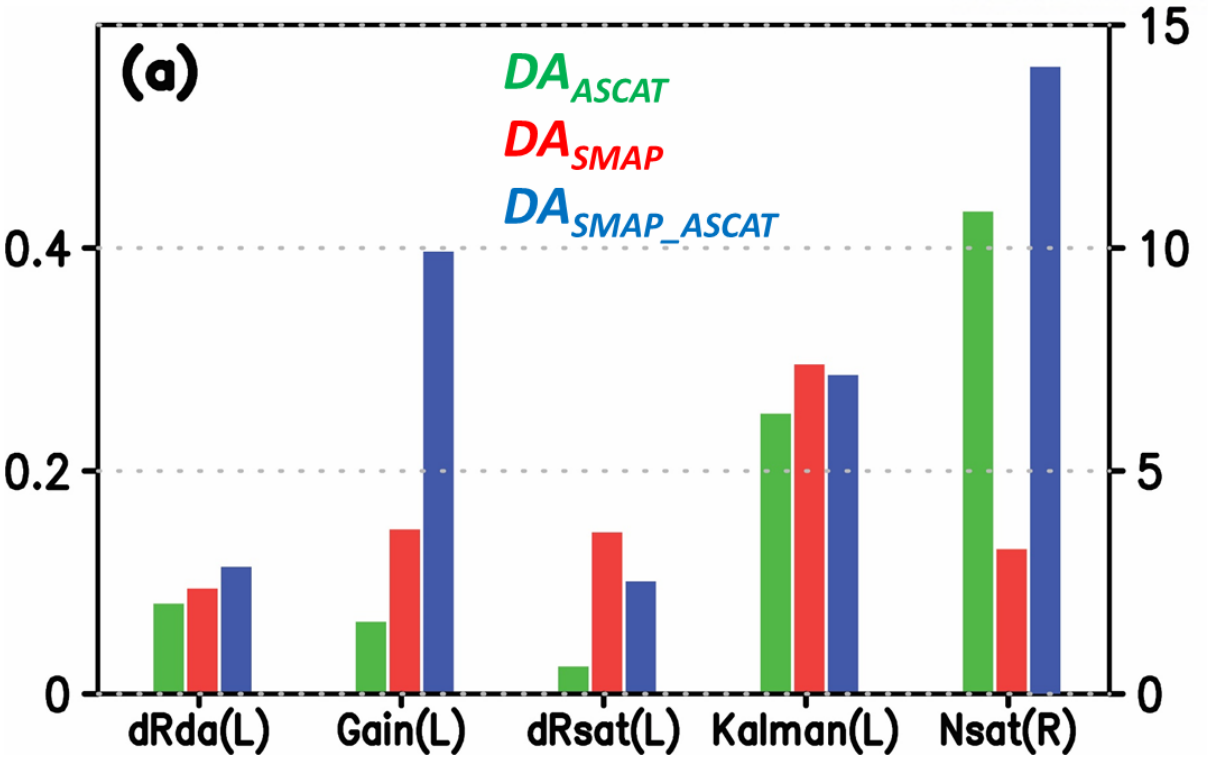

Figure 3.11 Mean values of soil moisture skill difference and components of the assimilation gain from the single-sensor assimilation of ASCAT (green) and SMAP (red) and the combined assimilation (blue) over the North America.

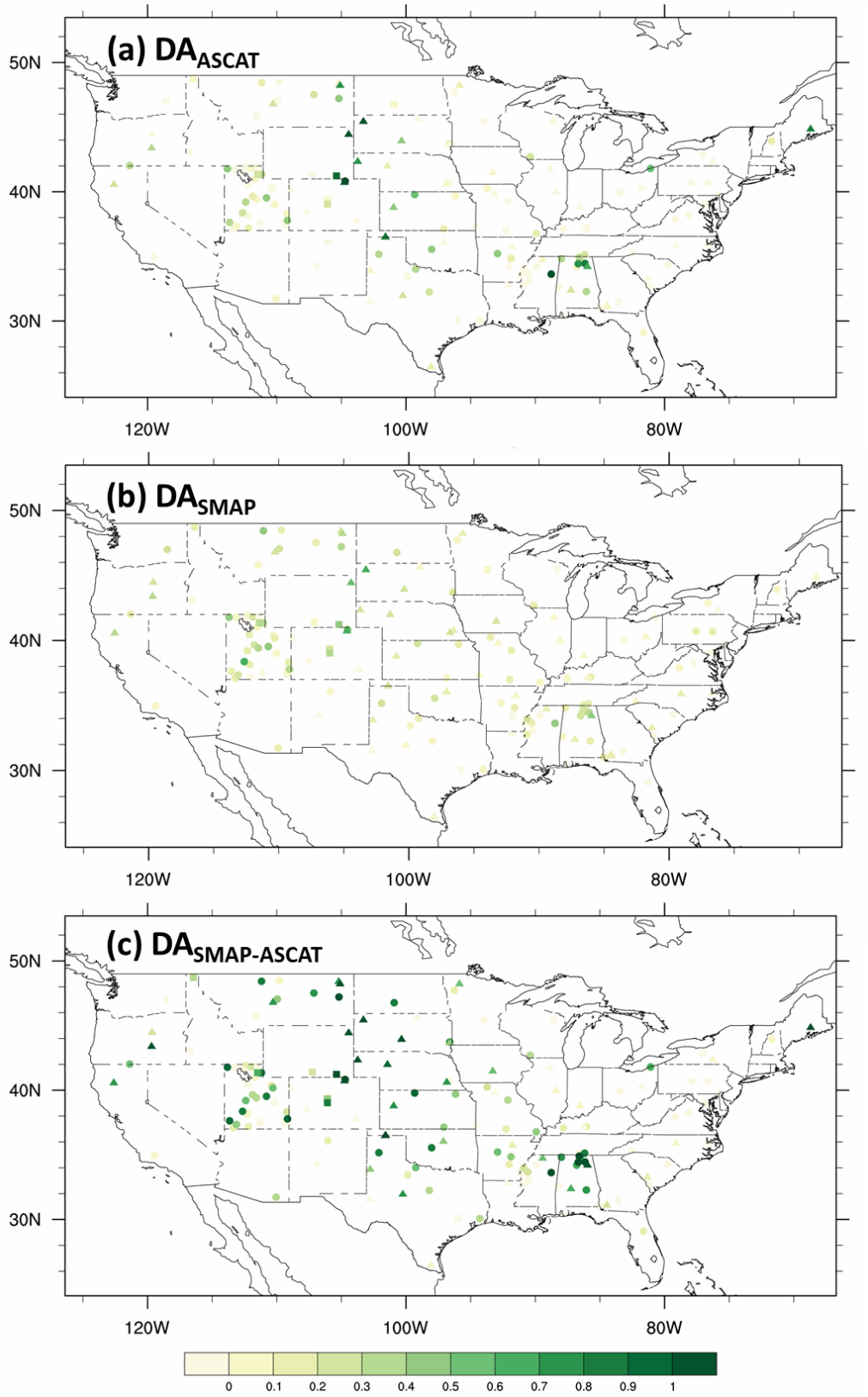

Figure 3.12 The spatial distribution of the assimilation gain from (a) ASCAT, (b) SMAP, and (c) joint retrieval assimilation experiments, evaluated by ground based in situ stations.

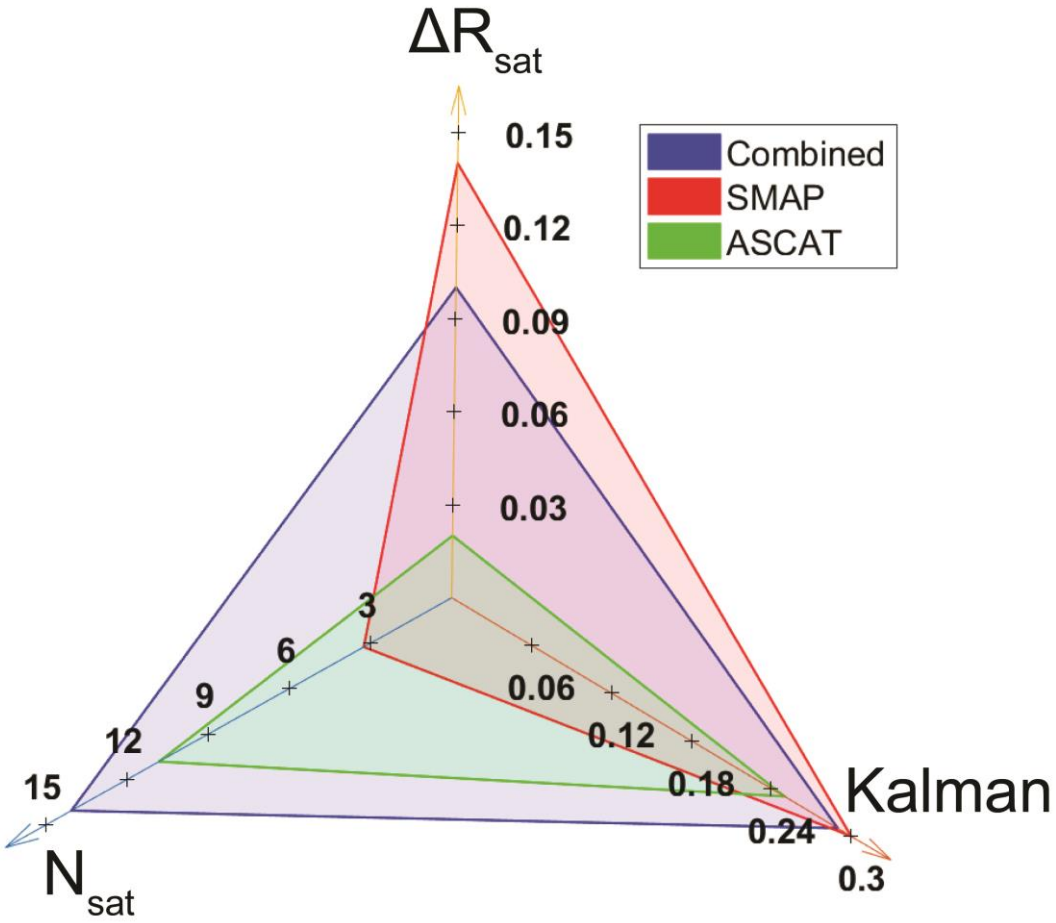

Figure 3.13 Radar chart of assimilation gain components (ΔR_{sat}: the soil moisture skill difference between satellite and the open loop model, Kalman: daily mean Kalman gain, and N_{sat}: The number of assimilated observation) from three different experiments (blue: ASCAT and SMAP combined assimilation, red: SMAP single-sensor assimilation, and green: ASCAT single-sensor assimilation) over the North America.

As this study perform the global soil moisture assimilation with the LETKF scheme, we can also evaluate the soil moisture estimates not only over the North America, but also over the other global regions such as the western Europe and the central Tibetan Plateau (Fig. 3.14). The global assimilation is a profit of the LETKF scheme which decomposes the global analysis domain into a number of independent sub-domains, even the computation resource is not much demanded. The satellite-based assimilation experiments consistently improve surface and root-zone soil moisture over the Europe and the Tibetan Plateau stations as well as over the North America. As consistent with the previous results, relatively larger improvements are obtained from the combined assimilation experiment (ASCAT and SMAP), where the skill improvement of surface and root-zone soil moisture over the Europe and the Tibetan Plateau is approximately 0.2 and 0.6, respectively. The soil moisture estimate is improved dramatically in the Tibet region because the performance of the open loop is significantly poor. The result is also reflected in the analysis averaged across categorized MODIS land cover as most of in situ sites in the Tibet region are classed in the grassland cover. Therefore, the assimilation shows that satellite retrievals provide significant added positive value not only for the North America, but also for the globe.

Figure 3.15 is relatively analogous to Figure 3.10 and confirm the main findings addressed in the previous analysis. That is, the difference between the combined (of ASCAT and SMAP) retrieval product skill and the open loop model is important to determine the skill improvement. The skill increase (ΔR) of the joint retrieval experiment relative to the open loop model, as a function of the R of the open loop model and of the assimilated (ASCAT and SMAP) observation. The skill improvement of surface and root-zone soil moisture through the data assimilation is strongly tided to the relative quality of the observed data compared to the open loop model.

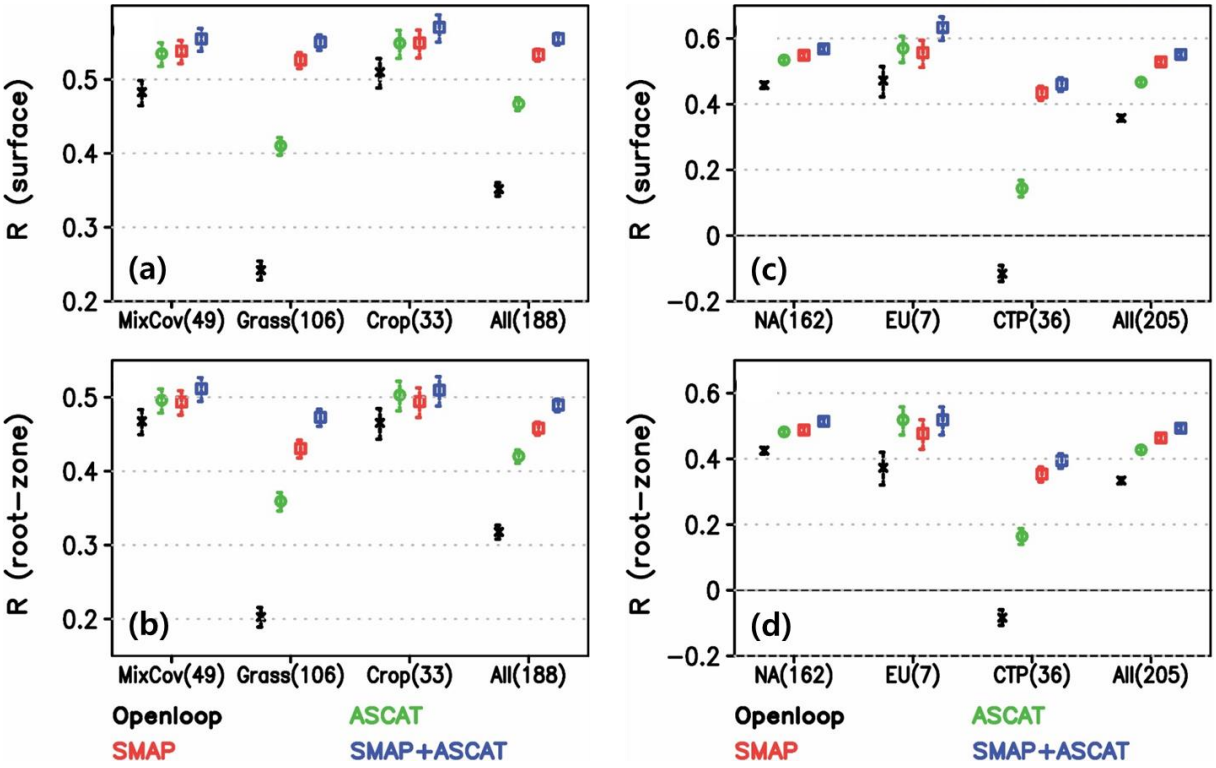

Figure 3.14 (a, c) Surface and (b, d) root-zone soil moisture skill from ASCAT (green), SMAP (red), and joint retrieval (blue) assimilation experiments with LETKF scheme and the open loop run, where errors represent 95% confidence intervals. The skills are averaged across categorized MODIS land cover (left column) and sub-regions of continents validated by in situ measurements over the globe.

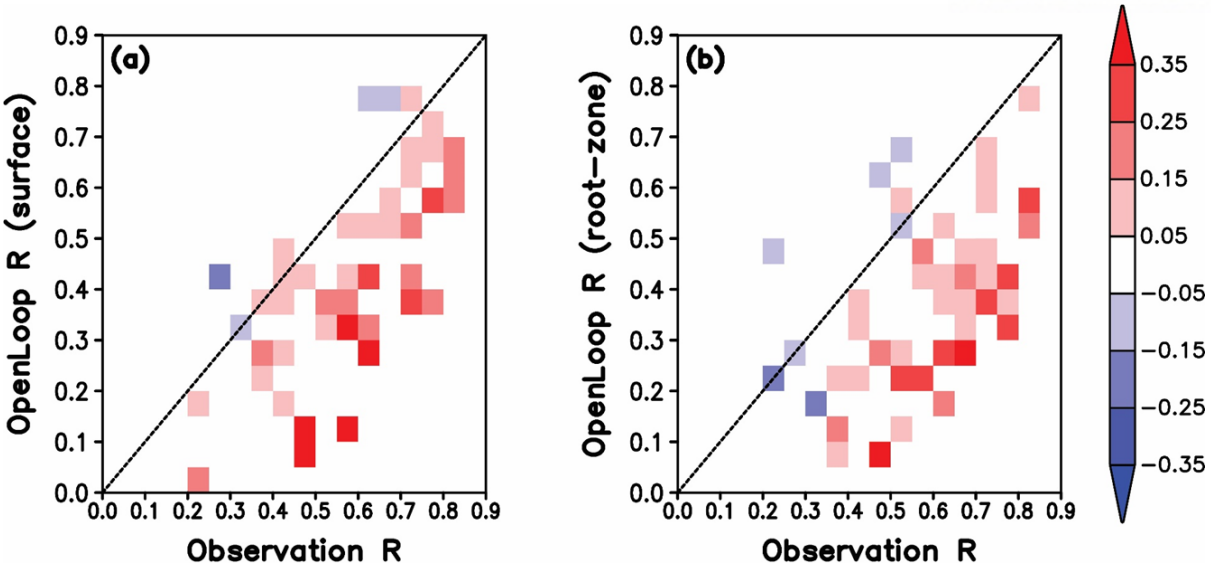

Figure 3.15 Skill improvement from the jointly assimilating both satellites of (a) surface and (b) root-zone soil moisture estimates, as a function of the open loop model skill (y-axis) and remotely sensed retrievals (x-axis). The shaded color represents the skill improvement defined as the skill of the assimilation product minus that of the open loop.

3.4.2. Dependency of active and passive remote sensing soil moisture retrievals

As shown in Figure 3.10, the skill of assimilated soil moisture estimates from all assimilation experiments is commonly increased, but the skill improvements are not much different even though active and passive remote sensing retrievals are separately used in each experiment. The spatial distribution of the skill improvement is also significant over the research domains, compared to the skill of soil moisture estimates simulated by the open loop (Fig. 3.16). However, there is regional dependency of skill improvement between ASCAT and SMAP assimilation experiment.

Figure 3.17 represents the spatial distribution of the skill difference of soil moisture estimates assimilated by SMAP and ASCAT satellites, evaluated by ground based in situ measurements. In the western North America, the soil moisture skill assimilated by SMAP tends to be better than those in the ASCAT assimilation, but in the central North America the results are almost opposite. In order to investigate the region dependency of the soil moisture skill difference, we separate the North America to region A (western North America: 120°W–105°W, 36°N–46°N) and B (central North America: 105°W–95°W, 26°N–47.5°N). There is positive anomaly ($\Delta R \approx 0.07$) of the skill difference map over the region A and its negative anomaly ($\Delta R \approx -0.04$) is shown in the region B, where the positive anomaly denotes that the quality of the soil moisture estimates assimilated with SMAP is better than the result of ASCAT, and vice versa.

As introduced in Section 2.5.2, we adopt the assimilation gain to identify which factor leads to these results. In the case of the region A, the assimilation gain of the SMAP assimilation is higher than those in the ASCAT assimilation corresponding to the skill difference between these assimilation experiments (Fig. 3.18). In particular, the relative skill of SMAP retrieval is most important factor to represent the high value of the assimilation gain rather than the Kalman gain and the number of sample (Fig. 3.19a). For the region B, the assimilation gain of the ASCAT assimilation is higher than those in the SMAP assimilation. The result is consistent with the better soil moisture skill of the ASCAT assimilation. Even though the relative skill and the Kalman gain of the SMAP experiment shows slightly higher than those of the ASCAT assimilation, the number of assimilated observation sample by ASCAT observations overwhelms that of SMAP (Fig. 3.19b). The reason why the number of assimilated observation samples of ASCAT is basically larger than that of SMAP is due to the difference of horizontal resolution of these satellites. However, the large difference of sample numbers is decreased over the region A is due to the quality control process for the remote sensing retrievals. Especially, ASCAT observations are discarded in which topographically complex provided by the satellite is larger than 10% and the complexity of the region A is large due to mountain terrain. Radar plots well explain that the assimilation gain is determined by the relative soil moisture skill of satellite datasets over the region A when the number of

assimilated observation is relatively comparable and the assimilation gain value is largely determined by the number of assimilated sample in the region B (Fig. 3.20).

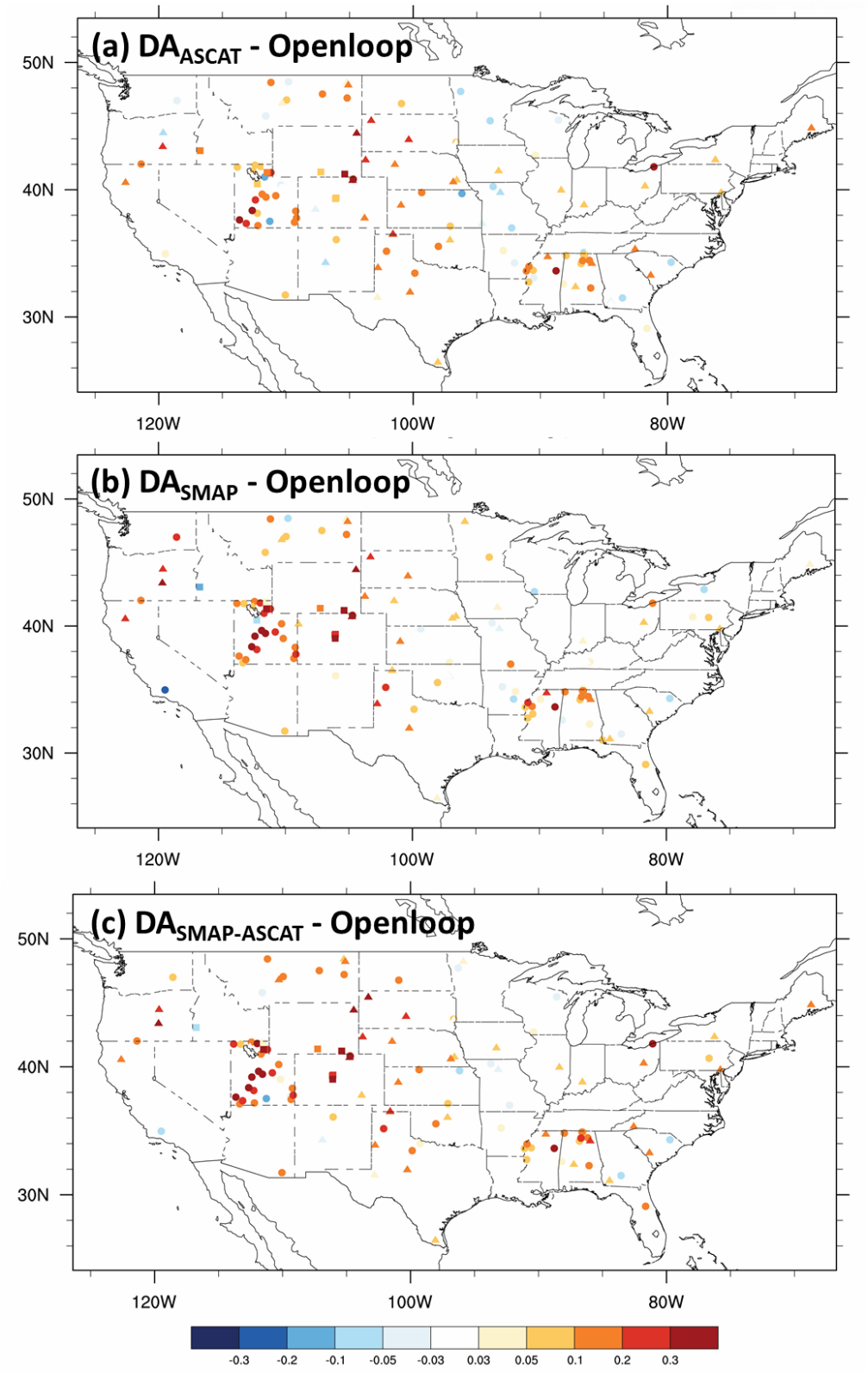

Figure 3.16 The spatial distribution of skill difference between assimilation experiments (a: ASCAT assimilation, b: SMAP assimilation, and c: joint retrieval assimilation) and the open loop simulation, evaluated by ground based in situ stations. Red color indicates a higher skill of the satellite assimilation

experiments and blue color indicates a higher skill of the open loop run.

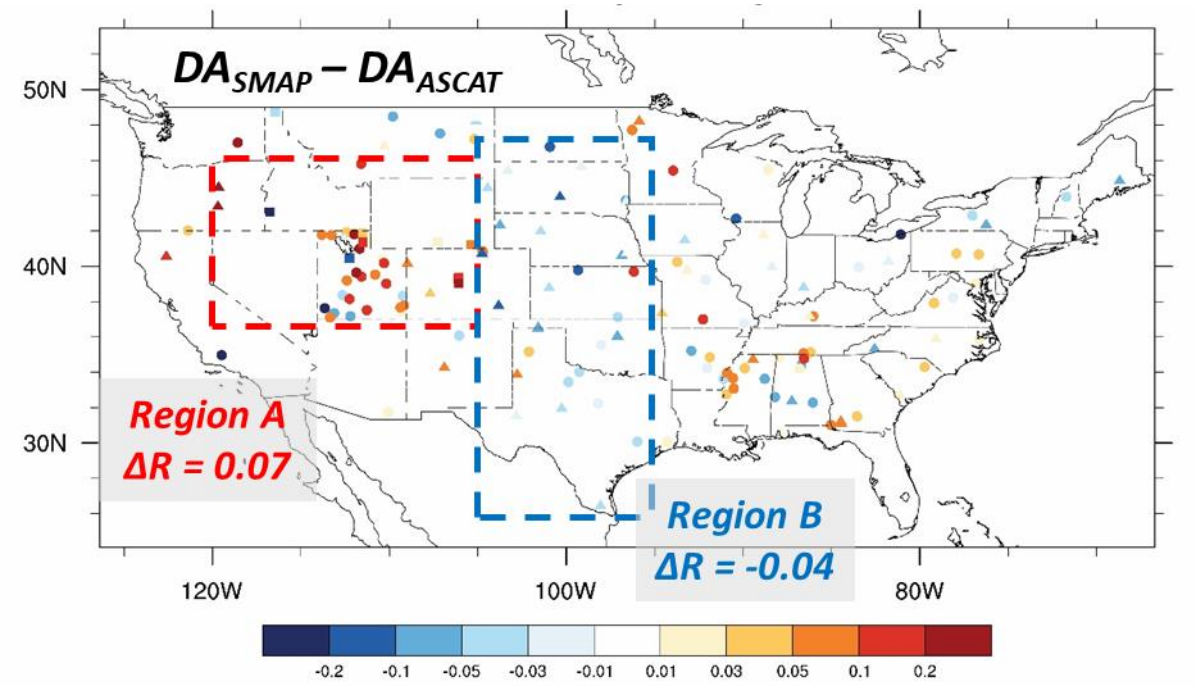

Figure 3.17 The spatial distribution of skill difference between the SMAP assimilation and the ASCAT assimilation, evaluated by ground based in situ stations. Red color indicates a higher skill of the SMAP assimilation experiment and blue color indicates a higher skill of the ASCAT assimilation.

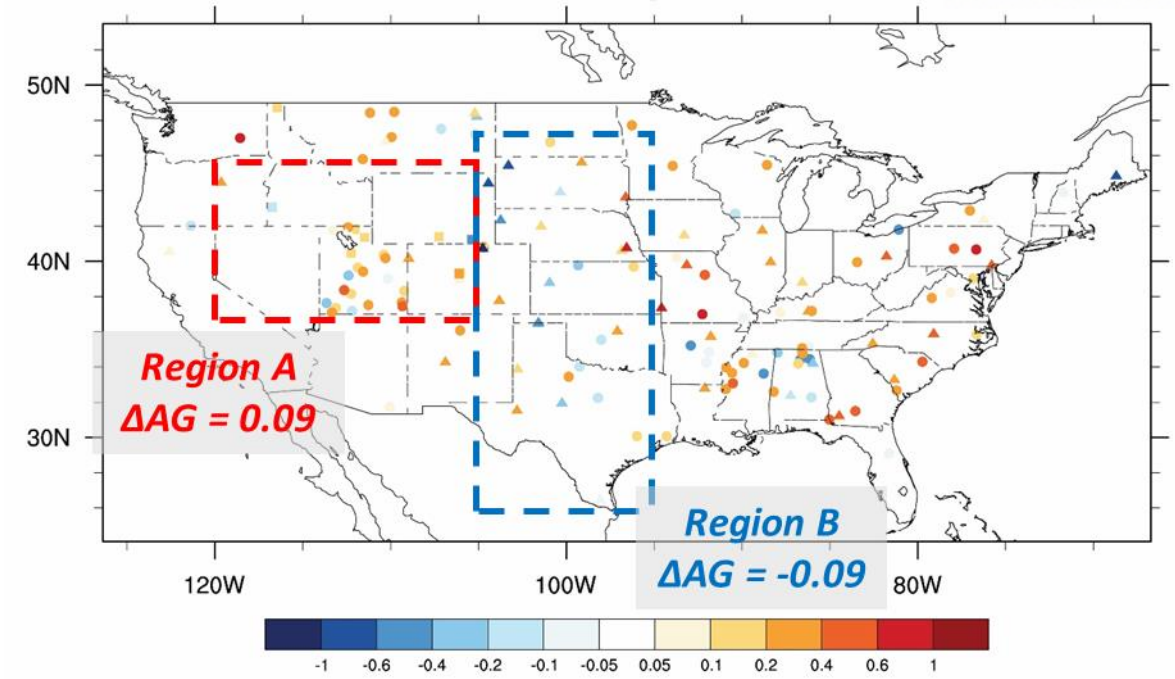

Figure 3.18 The spatial distribution of assimilation gain difference between the SMAP assimilation and the ASCAT assimilation, evaluated by ground based in situ stations. Red color indicates a higher impact of SMAP retrieval on the assimilation and blue color indicates a higher impact of ASCAT retrieval on the assimilation.

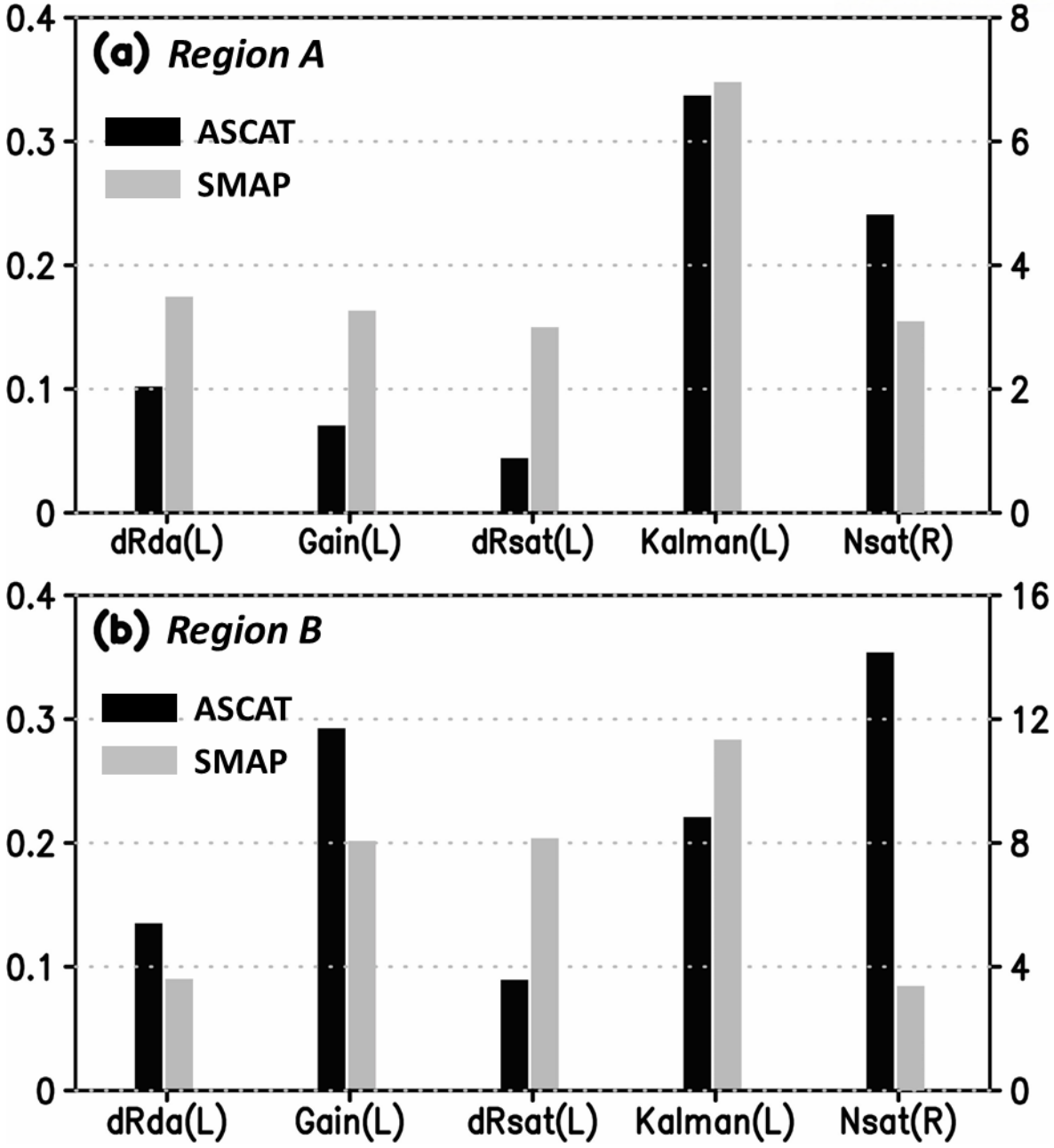

Figure 3.19 Mean values of assimilated soil moisture skill difference and components of the assimilation gain from SMAP (gray bar) and ASCAT (black bar) assimilation experiments over (a) region A (120°W–105°W, 36°N–46°N) and (b) region B (105°W–95°W, 26°N–47.5°N).

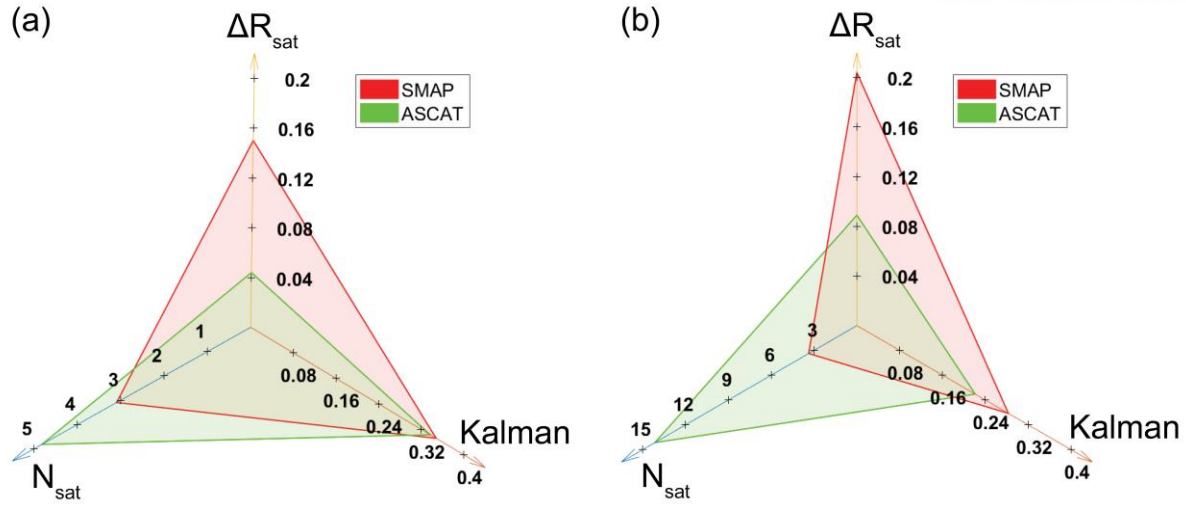

Figure 3.20 Radar chart of assimilation gain components (ΔR_{sat}: the soil moisture skill difference between satellite and the open loop model, Kalman: daily mean Kalman gain, and N_{sat}: The number of assimilated observation) from two single-sensor assimilation experiments (red: SMAP single-sensor assimilation, and green: ASCAT single-sensor assimilation) over (a) region A and (b) region B.

3.5. Implementation of assimilated soil moisture estimates

There is now emerging evidence that, soil moisture initial conditions could provide predictability at subseasonal to seasonal (S2S) time scale due to its persistency, at least in some regions of the world (R. D. Koster, Suarez, et al., 2004; Seneviratne, Koster, et al., 2006; Seo et al., 2018). As addressed in numerous previous studies, initialization of realistic land surface condition is important to improve the predictability of atmospheric variables as well as land surface conditions in dynamical forecast systems (R. Koster et al., 2011; R. D. Koster, Dirmeyer, et al., 2004; R. D. Koster et al., 2006; Seneviratne et al., 2010). Seo et al. (2018) highlighted that the impact of the soil moisture initialization on skill is found to be significant in the transitional region (June–August mean soil moisture climatology is 0.01–0.2 m^3 m^{-3}) between relatively dry and wet regions, and the improvement is sustained throughout the 2 months forecast period (Fig. 3.21). The lack of appreciable skill improvement in the dry or wet regimes is probably due to the relative insensitivity of evapotranspiration to soil moisture variation in that regimes. In other words, when we prescribe the realistic land conditions in dynamical forecast models, it leads to a better prediction of the atmospheric states in the transitional region through the realistic representation of land-atmosphere interaction. Therefore, it is necessary to produce the best analysis of land surface states, especially, over the transitional region in order to improve the better predictability through the land initialization in the model.

In this section, this study categorizes the skill improvement of soil moisture estimates by the data assimilation according to background soil moisture climatology over the North America for the research period (each category covers approximately 10% of total in situ sites over the North America region). Figure 3.22 reveals the mean skill of surface and root-zone soil moisture estimates from the joint assimilation with ASCAT and SMAP satellite datasets and the open loop run categorized by soil moisture climatology for the research period. The skill improvement through the data assimilation occurs primarily in soil moisture climatology of 0.11–0.25 m^3/m^3, where the skill difference of surface and root-zone soil moisture between the assimilation and the open loop is approximately 0.19 and 0.16, respectively. Hence, if the assimilated soil moisture is initialized in the forecast system, we can expect the better representation of atmospheric states in the transitional regions.

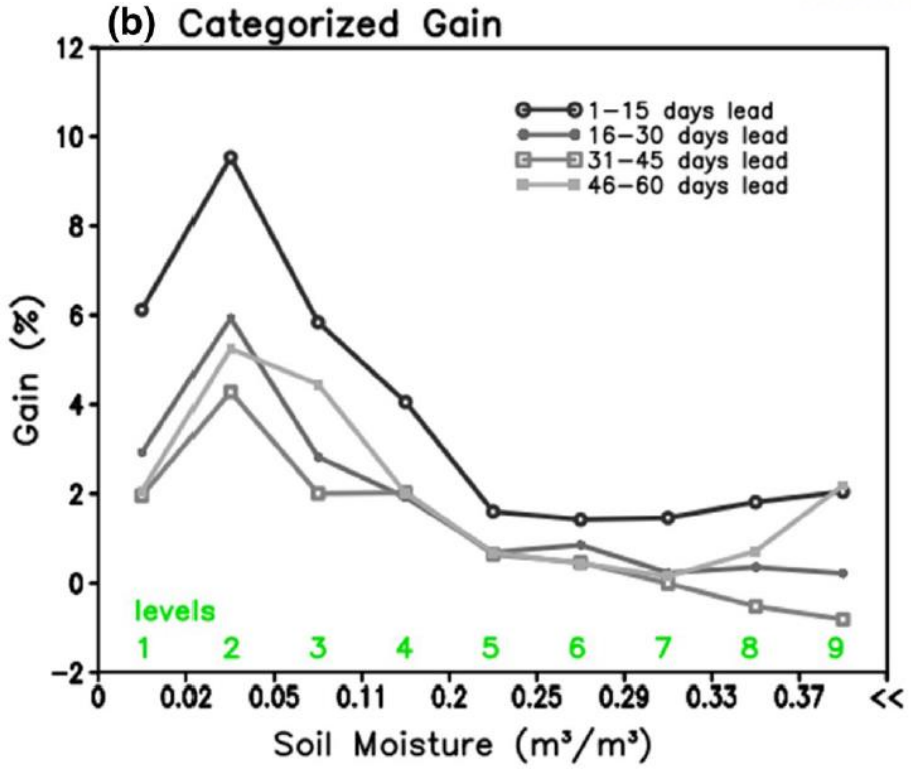

Figure 3.21 Gain values of surface air temperature by soil moisture initialization in AMIP-type simulation of GloSea5 for each lead time as a function of JJA mean soil moisture climatology for 1996–2010. This figure is Fig. 6b of (Seo et al., 2018).

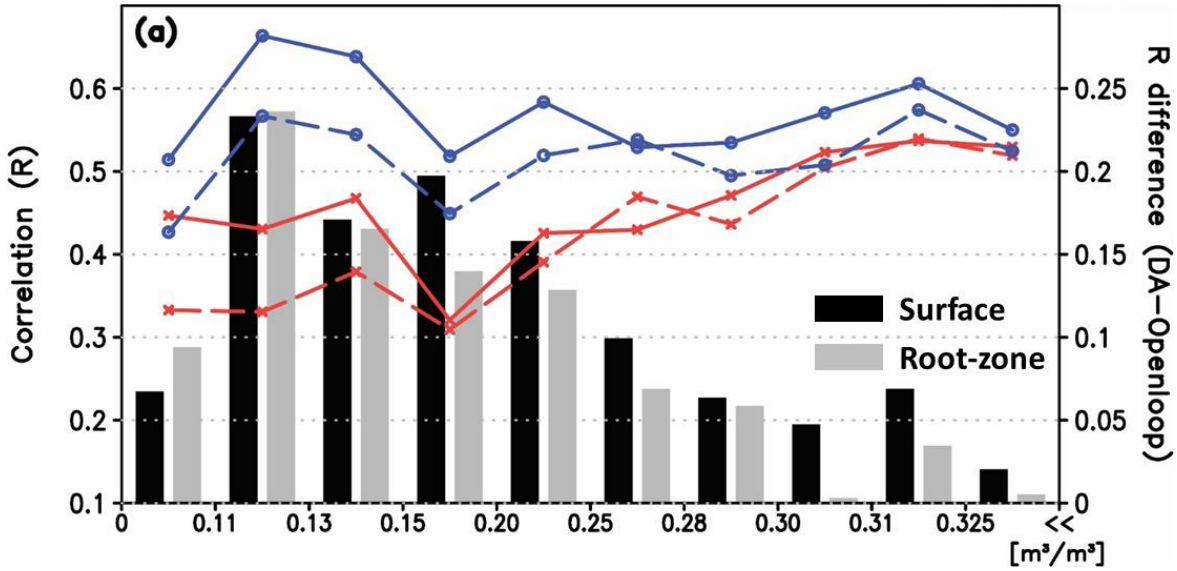

Figure 3.22 Mean skill of surface (solid lines) and root-zone (dashed lines) soil moisture from the soil moisture assimilation (blue) and the open loop run (red), averaged across categorized soil moisture climatology for the research period. Bar graphs represent the skill difference between the assimilation experiment and the open loop, where black and gray bar denote the difference of surface and root-zone soil moisture estimates, respectively.

Figure 3.23 represents the U.S. drought monitoring for drought developed period. The drought signals are intensified and extended, so that the all of states located at west coast of North America are suffered by the severe drought in California, Nevada, Utah, Arizona, Alabama, Georgian, and Tennessee states. The classification of drought intensity is based on five key indicators along with several other objective indicators. The five indicators consist of Palmar Drought Severity Index (PDSI), CPC soil moisture model, USGS weekly streamflow, Standardized Precipitation Index (SPI), and objective drought indicator blends. Since two of them include soil moisture variable, improved soil moisture information through data assimilation can be important information for realistic drought monitoring.

Figure 3.24 represents the spatial distribution of drought index based on surface soil moisture. As described in Section 2.5.3, the value closed to zero means is the evidence of extreme drought. PCI and TCI based on meteorology variables do not represent the mesoscale regional locality, but the land surface condition enables to resolve the regional locality. Therefore, to understand the drought phenomenon, the realistic land surface condition is as important as the meteorology variables. Furthermore, when the soil moisture derived by the model is assimilated with the satellite datasets, the spatial distribution of land surface dryness tends to represent better than those in the open loop model. For instance, the drought monitoring over the U.S. continent for the drought active period indicates that California, Nevada, Arizona and Utah states are suffered by the extreme drought. The observed extreme climate event is realistically revealed in the assimilated soil moisture state.

When we look at the daily time series of surface and root-zone soil moisture from in situ observation, combined assimilation experiment (Fig. 3.25), and the open loop run over Utah state, the relative dryness is well captured in the assimilation experiment for surface and root-zone layers during 20 May–20 June 2016. The result from the assimilation experiment shows the observed surface soil moisture of 0.1 m^3/m^3 or less although it is not possible to simulate the observed dryness in the root-zone layer.

Corresponding the realistic daily variation of soil moisture, the improvement of the land-atmosphere interaction is also expected through the soil moisture data assimilation. As described in daily evolution of the soil moisture estimates over the Utah, there is abnormal drought for 1 month from the last of May 2016. As described in Eq. 2.29, the soil moisture-temperature coupling from the assimilation experiment appears strong for to the corresponding drought period, but the strong interaction is not represented in the open loop model (Fig. 3.26). The improvement of the land-atmosphere interaction by the soil moisture assimilation is significant in the time series of its difference, where the abnormal coupling is well captured by the 2016 drought event.

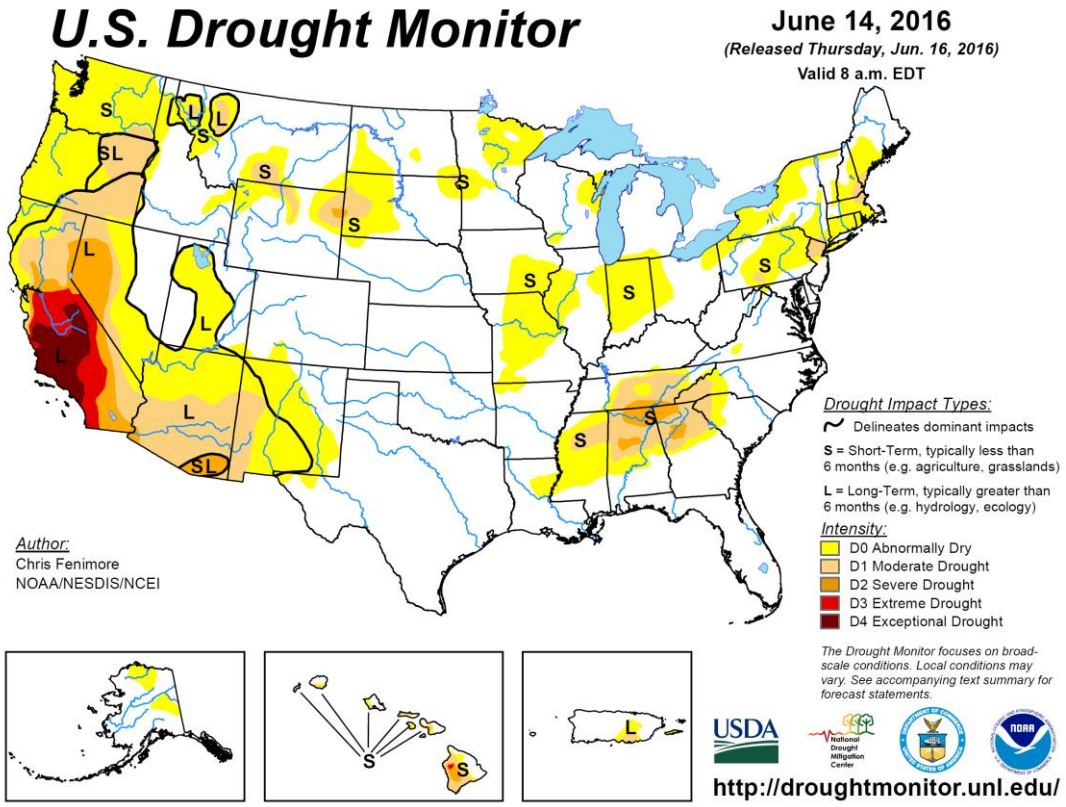

Figure 3.23 Map of U.S. drought monitoring during 14 June 2016. (http://droughtmonitor.unl.edu)

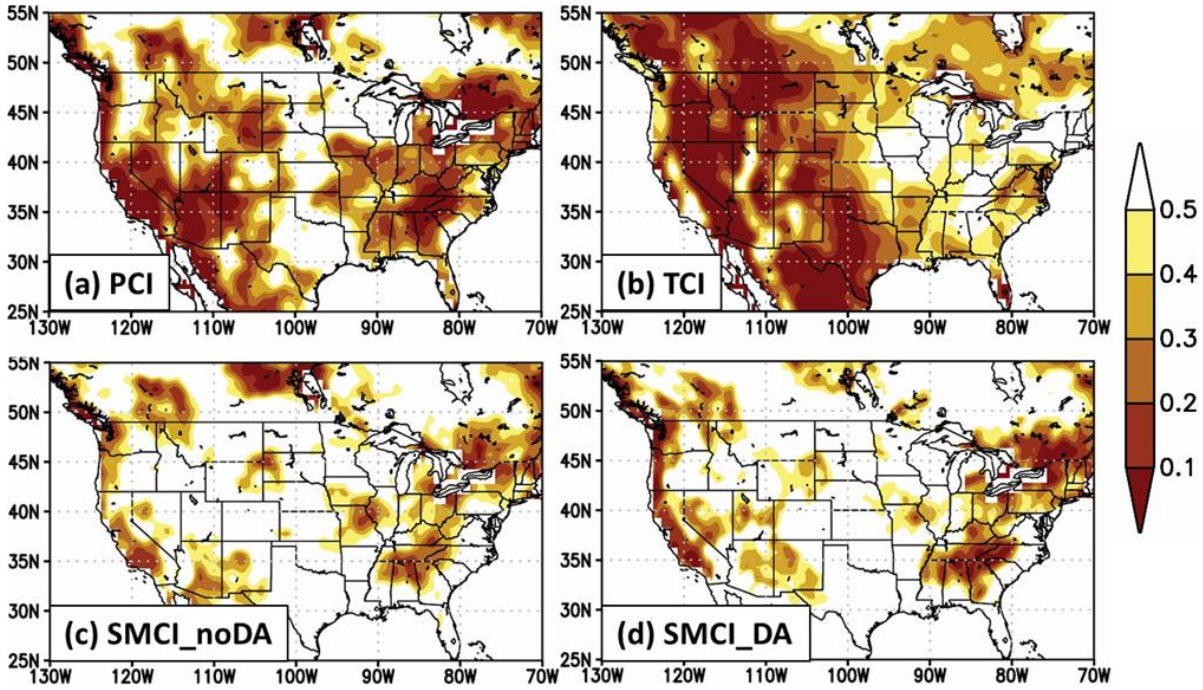

Figure 3.24 The spatial distribution of drought indices of (a) PCI, (b) TCI, (c) SMCI from the open loop model, and (d) SMCI from the assimilation experiment during 20 May–20 June 2016.

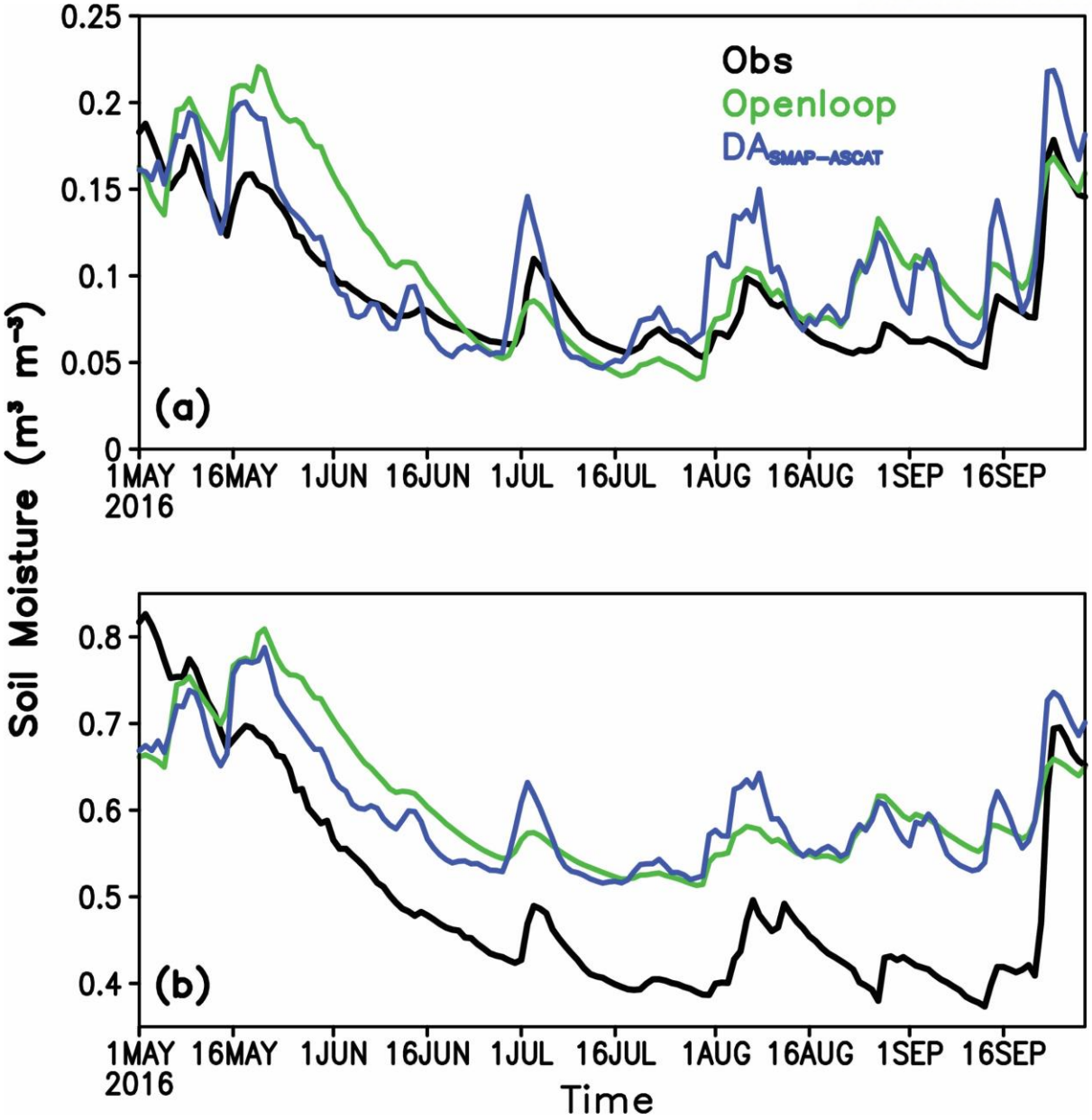

Figure 3.25 Daily time series of (a) surface and (b) root-zone soil moisture from in situ observation (black), combined assimilation experiment (blue), and the open loop run (green) over Utah state (115.5°W–107.5°W, 36°N–43°N). The mean of 36 number of in situ observation are used in the validation.

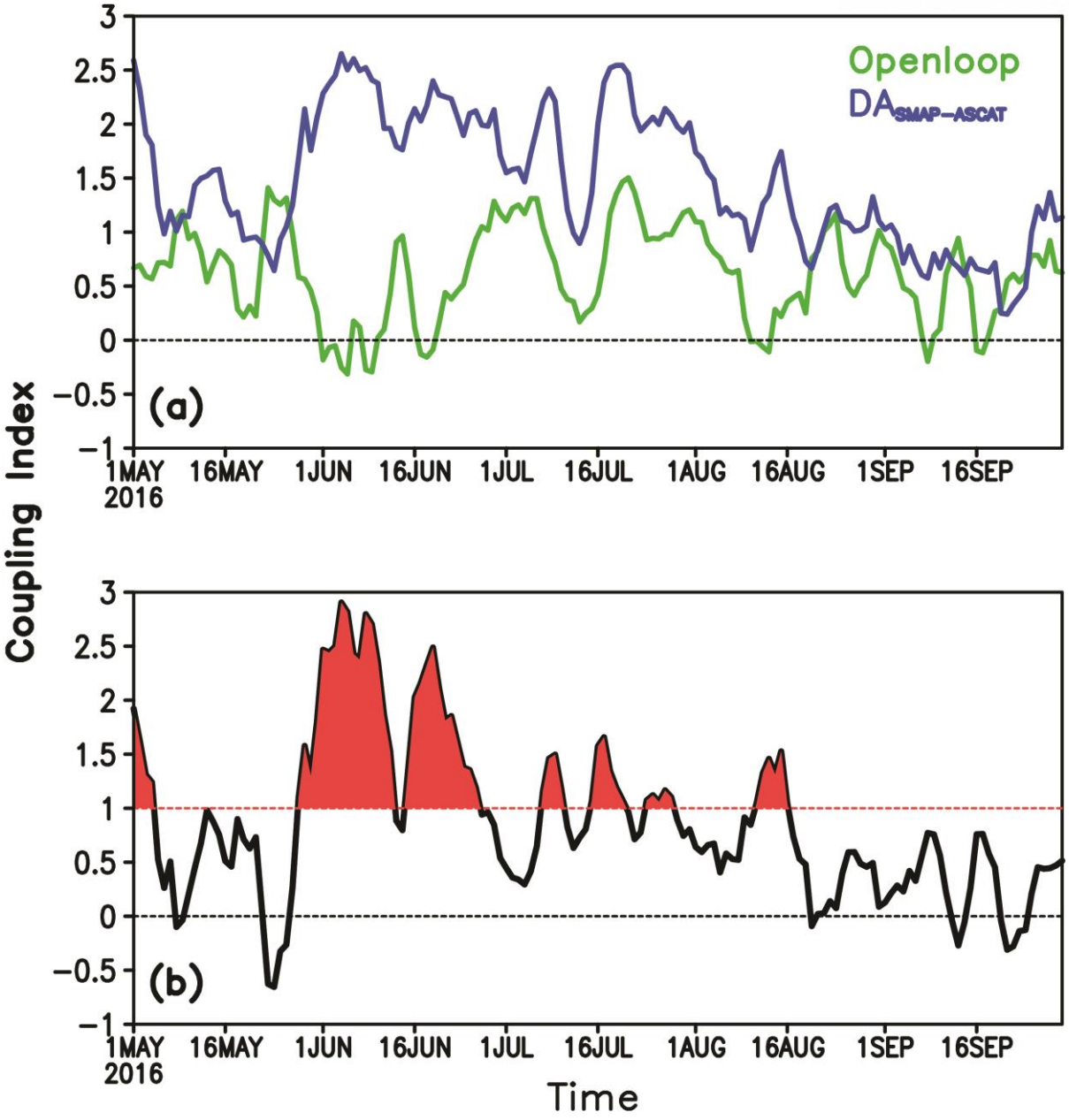

Figure 3.26 Daily time series of (a) soil moisture-temperature coupling index from combined assimilation experiment (blue), and the open loop run (green) and (b) its difference between them in which red shading represents the strong land-atmosphere coupling over Utah state (115.5°W–107.5°W, 36°N–43°N).

3.6. Summary

In this section, we establish the framework of optimized soil moisture data assimilation system based on LETKF scheme and document the impact of assimilation schemes and remote sensing retrievals on the assimilated soil moisture estimates.

In order to operate the global data assimilation system with efficient computation resource, this study adopt LETKF scheme. In this method, several parameters significantly influence on the assimilation process such as the localization scale parameter and the covariance inflation parameter. For the localization parameter, as increasing the parameter, which means that large distance observations are innovated in the assimilation, the skill improvement of surface and root-zone soil moisture becomes marginal because the uncorrelated distance observations mostly provide negative value. On the other hand, the inflation parameter shows the different sensitivity depending on the assimilated satellite dataset. For instance, in the case of the SMAP data, the assimilated soil moisture is improved as the parameter is increased, while the ASCAT sensitive experiments show the opposite result. Hence the localization scale parameter is determined as narrow as possible, and the covariance inflation parameter used the intermediate value considering the results of each satellite.

Based on this data assimilation framework, we investigate the impact of assimilation schemes and remote sensing retrievals on the assimilated soil moisture estimates. The results of sensitivity tests according to assimilation schemes suggest that a considerable computation efficiency is shown in the assimilation based on the LETKF scheme. In terms of the soil moisture skill, the result by the efficient scheme shows comparable or better skill to the conventional EnKF method. Moreover, this study conducts three set of soil moisture assimilation experiments of the single-sensor assimilation experiments (of ASCAT or SMAP) and the combined assimilation experiment (ASCAT and SMAP) with the LETKF scheme. The result of the assimilation experiments shows that both of satellite retrievals provide a significant positive value on the model and the skill improvement of surface and root-zone soil moisture through the assimilating active and passive microwave satellites is approximately up to ~0.1. The influence of the data assimilation is strongly tided to the skill of the open loop model. The skill improvement of surface and root-zone soil moisture through the data assimilation is significant as the relative quality of the observed data is much better than that of the open loop model. The soil moisture skill is mostly improved in the grassland class where the performance of the open loop is worst among the other land cover classes.

The results are tried to be explained by a newly introduced “Assimilation Gain” and its components. The joint assimilation experiment performs best, which refers to the complementary value of active and passive remoted sensed datasets, and the single-sensor assimilation experiments yield similar

improvements, and the improvement is highly tided to the relative skill of assimilated satellite data to the open loop model influences. However, the skill difference between the single-sensor assimilation experiments shows regional dependency according to assimilated satellite dataset. For instance, the soil moisture from the SMAP single-sensor assimilation in the western North America shows better performance than those in the ASCAT and the result is opposite in the central North America. By the specific demonstration, we found that the relative skill of assimilated satellite data to the open loop model influences on the western North America and the number of assimilated observation sample is related to the result over the central North America. In other words, if the number of observed samples is same, the results are determined by the quality of the observations and, in the opposite case, if the quality of the observations is comparable, the number of observations affects the results.

In addition, the result of improving soil moisture estimates through the data assimilation is related to the background state of soil moisture climatology. The improvement is dominantly concentrated on the transitional region where the predictability of atmospheric variables is improved when the land-atmosphere interaction process is realistically represented in the dynamical forecast model. It means that a better predictive performance can be expected in the forecast model if the land-atmosphere process is to be realistic by initializing the improved soil moisture data through the assimilation, especially, over the transitional area.

Chapter 4

4. Conclusion

This study suggests the realistic land reanalysis data is needed to understand the precise land-atmosphere interactions. The interest of understanding land-atmosphere interaction has been attended since its impact directly affect our lives and will be intensified in the future climate. The qualified soil moisture data is determined by: (1) LSM physics, (2) atmospheric boundary forcing variables, and (3) soil moisture data assimilation process. The impact of the first and second contributors has been deeply investigated by the Land Data Assimilation System (LDAS) projects. Although the number of available observations has been increased recently, the third contributor is not much researched. Sequentially, the soil moisture data assimilation is influenced by the assimilation schemes and the remote sensing retrievals, so that this study mainly focuses on these impacts on the producing realistic land surface conditions.

This study uses a data assimilation system based on the LETKF technique which is optimized through several sensitivity experiments (e.g. localization scale parameter and covariance inflation parameter). The system assimilates C-band active (ASCAT) and L-band passive (SMAP) microwave remote sensing soil moisture observations, but the retrievals are subjected to quality control and bias correction process based on cumulative distribution function fitting. Based on this data assimilation framework, we examine the impact of assimilation schemes and remote sensing retrievals on the assimilated soil moisture estimates and evaluate the soil moisture estimates with ground based in situ measurements. We compared the results from EnKF and LETKF assimilation experiments in the data assimilation process and confirmed that the LETKF scheme is a more efficient method in terms of computing time than the EnKF without significantly soil moisture skill decrease. In addition, this study also conducts three set of soil moisture assimilation experiments of the single-sensor assimilation experiments (either ASCAT or SMAP) and the combined assimilation experiment (ASCAT and SMAP) with the LETKF scheme in order to evaluate the impact of each satellite on the assimilation results. The joint assimilation shows the best performance in which surface and root-zone soil moisture skill improvement is about 0.12 and the skill increase from the single-sensor assimilation experiments is approximately up to ~0.1, when compared to the open loop model. The skill improvement of surface and root-zone soil moisture from the three experiments is mostly significant in the grassland class because the soil moisture skill from the open loop has significantly lower over the land cover. On the other hand, the soil moisture skill difference between the single-sensor assimilation experiments reveals

the regional dependency. The result enables to be explained by “Assimilation Gain” which is composed by the relative retrieval skill, the Kalman gain, and the number of assimilated observational sample.

The soil moisture skill increase from the data assimilation is seen in the transitional regions in which the relative sensitivity of evapotranspiration to soil moisture variations is high. Therefore, if we initialize the realistic land conditions in dynamical forecast models, it leads to a better prediction of the atmospheric states in the transitional region through the realistic representation of land-atmosphere interaction. Actually, the assimilated soil moisture tends to represent better other radiational variables as well as soil moisture skill, which results in the realistic soil moisture-temperature coupling. It is very promising to enhance the predictability at S2S time scale that soil moisture states have a memory.

Chapter 5

5. Outlook and future works

The interest in surface and subsurface soil moisture has been attended by their important role in the partitioning of the available energy incident at the land surface into sensible and latent heat fluxes. The land-atmosphere interaction influences on the local atmospheric processes such as the boundary layer height and cloud coverage (Betts & Ball, 1998) and leads to a robust circumglobal teleconnection response across the planetary length scale (Teng et al., 2019). The interaction is strongly tied to develop extreme climate events (e.g. mega heatwave, flood and drought) in recent global warming (Perkins, Alexander, & Nairn, 2012) and will be reinforced in the projected future climate (Dirmeyer et al., 2013). Furthermore, subsurface soil moisture exhibits a persistency on weekly to monthly time scales (R. D. Koster & Suarez, 2001; Seneviratne, Koster, et al., 2006), so that the realistic soil moisture estimates is useful to initialize the land surface states. It may lead to enhance the forecast skill of atmospheric status in sub-seasonal to seasonal time scale, where the quality of initialized soil moisture estimates determines the improvement of the predictability.

In this study, we just focus on 2016 due to the limitation of computation and data availability. However, in recent decade, there are other available soil moisture retrievals such as ASCAT, SMOS, AMSR2, and SMAP from 2007 summer to present. Therefore, I will produce long-term assimilated soil moisture estimates with the entire satellite datasets. Based on the land surface reanalysis, I will examine the impact of the qualified soil moisture initialization on S2S time scale forecasts. The realistic initial condition significantly enhances the forecast skill in the dynamical prediction systems through the representation of realistic land-atmosphere interaction. UK Met Office Global Seasonal forecast system version 5 (GloSea5) will be used to perform two sets of hindcast experiments depending on whether realistic soil moisture is initialized or not. Therefore, the qualified land reanalysis is widely used in the understanding land surface process as well as the initialization of realistic land surface boundary conditions.

References

Albergel, C., De Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., . . . Wagner, W. (2012). Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment, 118*, 215-226.

Andreadis, K. M., & Lettenmaier, D. P. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources, 29*(6), 872-886.

Aonashi, K., Awaka, J., Hirose, M., Kozu, T., Kubota, T., Liu, G., . . . Takahashi, N. (2009). GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation. *Journal of the Meteorological Society of Japan. Ser. II, 87*, 119-136.

Best, M., Pryor, M., Clark, D., Rooney, G., Essery, R., Ménard, C., . . . Gedney, N. (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. *Geoscientific Model Development, 4*(3), 677-699.

Betts, A. K., & Ball, J. H. (1998). FIFE surface climate and site-average dataset 1987–89. *Journal of the Atmospheric Sciences, 55*(7), 1091-1108.

Bosilovich, M. G., Radakovich, J. D., da SILVA, A., Todling, R., & Verter, F. (2007). Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *Journal of the Meteorological Society of Japan. Ser. II, 85*, 205-228.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., . . . Llorens, P. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sensing of Environment, 115*(12), 3390-3408.

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., . . . Piepmeier, J. (2016). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing, 54*(8), 4994-5007.

Cho, E., Su, C.-H., Ryu, D., Kim, H., & Choi, M. (2017). Does AMSR2 produce better soil moisture retrievals than AMSR-E over Australia? *Remote Sensing of Environment, 188*, 95-105.

Conil, S., Douville, H., & Tyteca, S. (2007). The relative influence of soil moisture and SST in climate predictability explored within ensembles of AMIP type experiments. *Climate dynamics, 28*(2-3), 125-145.

Crow, W. T., & Wood, E. F. (2003). The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Advances in Water Resources, 26*(2), 137-149.

Dirmeyer, P. A. (2003). The role of the land surface background state in climate predictability. *Journal of hydrometeorology, 4*(3), 599-610.

Dirmeyer, P. A. (2005). The land surface contribution to the potential predictability of boreal summer

season climate. *Journal of hydrometeorology, 6*(5), 618-632.

Dirmeyer, P. A., Jin, Y., Singh, B., & Yan, X. (2013). Trends in land–atmosphere interactions from CMIP5 simulations. *Journal of hydrometeorology, 14*(3), 829-849.

Dobson, M. C., & Ulaby, F. T. (1986). Active microwave soil moisture research. *IEEE Transactions on Geoscience and Remote Sensing*(1), 23-36.

Dorigo, W., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., De Jeu, R. A., & Naeimi, V. (2010). Error characterisation of global active and passive microwave soil moisture data sets.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., . . . Gruber, A. (2017). ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art and future directions. *Remote Sensing of Environment, 203*, 185-215.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., . . . Oevelen, P. v. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences, 15*(5), 1675-1698.

Douville, H. (2004). Relevance of soil moisture for seasonal atmospheric predictions: is it an initial value problem? *Climate dynamics, 22*(4), 429-446.

Douville, H., & Chauvin, F. (2000). Relevance of soil moisture for seasonal climate predictions: A preliminary study. *Climate dynamics, 16*(10-11), 719-736.

Draper, C., Reichle, R., De Lannoy, G., & Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters, 39*(4).

Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. *Journal of Geophysical Research: Atmospheres, 112*(D3).

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans, 99*(C5), 10143-10162.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics, 53*(4), 343-367.

Evensen, G. (2009). *Data assimilation: the ensemble Kalman filter*: Springer Science & Business Media.

Ferranti, L., & Viterbo, P. (2006). The European summer of 2003: Sensitivity to soil water initial conditions. *Journal of Climate, 19*(15), 3659-3680.

Fischer, E. M., Seneviratne, S., Vidale, P., Lüthi, D., & Schär, C. (2007). Soil moisture–atmosphere interactions during the 2003 European summer heat wave. *Journal of Climate, 20*(20), 5081-5099.

Ford, T. W., Dirmeyer, P. A., & Benson, D. O. (2018). Evaluation of heat wave forecasts seamlessly across subseasonal timescales. *npj Clim. Atmos. Sci., 1*, 20.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment, 114*(1), 168-182.

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing, 55*(12), 6780-6792.

Guo, Z., Dirmeyer, P. A., & DelSole, T. (2011). Land surface impacts on subseasonal and seasonal predictability. *Geophysical Research Letters, 38*(24).

Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review, 129*(11), 2776-2790.

Hauser, M., Orth, R., & Seneviratne, S. I. (2016). Role of soil moisture versus recent climate change for the 2010 heat wave in western Russia. *Geophysical Research Letters, 43*(6), 2819-2826.

Hoskins, B., Fonseca, R., Blackburn, M., & Jung, T. (2012). Relaxing the tropics to an 'observed'state: Analysis using a simple baroclinic model. *Quarterly Journal of the Royal Meteorological Society, 138*(667), 1618-1626.

Houtekamer, P., & Zhang, F. (2016). Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review, 144*(12), 4489-4532.

Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review, 129*(1), 123-137.

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena, 230*(1-2), 112-126.

Jeong, J.-H., Linderholm, H. W., Woo, S.-H., Folland, C., Kim, B.-M., Kim, S.-J., & Chen, D. (2013). Impacts of snow initialization on subseasonal forecasts of surface air temperature for the cold season. *Journal of Climate, 26*(6), 1956-1972.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., . . . Endo, H. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II, 93*(1), 5-48.

Kolassa, J., Reichle, R., & Draper, C. S. (2017). Merging active and passive microwave observations in soil moisture data assimilation. *Remote Sensing of Environment, 191*, 117-130.

Koster, R., Mahanama, S., Yamada, T., Balsamo, G., Berg, A., Boisserie, M., . . . Gordon, C. (2011). The second phase of the global land–atmosphere coupling experiment: soil moisture contributions to subseasonal forecast skill. *Journal of hydrometeorology, 12*(5), 805-822.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., . . . Lawrence, D. (2004). Regions

of strong coupling between soil moisture and precipitation. *Science, 305*(5687), 1138-1140.

Koster, R. D., Dirmeyer, P. A., Hahmann, A. N., Ijpelaar, R., Tyahla, L., Cox, P., & Suarez, M. J. (2002). Comparing the degree of land–atmosphere interaction in four atmospheric general circulation models. *Journal of hydrometeorology, 3*(3), 363-375.

Koster, R. D., Mahanama, S., Yamada, T., Balsamo, G., Berg, A., Boisserie, M., . . . Gordon, C. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters, 37*(2).

Koster, R. D., & Suarez, M. J. (2001). Soil moisture memory in climate models. *Journal of hydrometeorology, 2*(6), 558-570.

Koster, R. D., Suarez, M. J., Liu, P., Jambor, U., Berg, A., Kistler, M., . . . Famiglietti, J. (2004). Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *Journal of hydrometeorology, 5*(6), 1049-1063.

Koster, R. D., Sud, Y., Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., . . . Davies, H. (2006). GLACE: the global land–atmosphere coupling experiment. Part I: overview. *Journal of hydrometeorology, 7*(4), 590-610.

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., . . . Nakagawa, K. (2007). Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Transactions on Geoscience and Remote Sensing, 45*(7), 2259-2275.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., . . . Dirmeyer, P. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environmental modelling & software, 21*(10), 1402-1415.

Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T., . . . Houser, P. R. (2008). A land surface data assimilation framework using the land information system: Description and applications. *Advances in Water Resources, 31*(11), 1419-1432.

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G., Dunbar, R. S., Kim, S., . . . Wagner, W. (2017). Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters, 44*(12), 6145-6153.

Liu, Q., Reichle, R. H., Bindlish, R., Cosh, M. H., Crow, W. T., de Jeu, R., . . . Jackson, T. J. (2011). The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *Journal of hydrometeorology, 19*(2).

Liu, Y., Dorigo, W. A., Parinussa, R., de Jeu, R. A., Wagner, W., McCabe, M. F., . . . Van Dijk, A. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment, 123*, 280-297.

Loveland, T. R., & Belward, A. (1997). The IGBP-DIS global 1km land cover data set, DISCover: first

results. *International Journal of Remote Sensing, 18*(15), 3289-3295.

Margulis, S. A., McLaughlin, D., Entekhabi, D., & Dunne, S. (2002). Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water resources research, 38*(12).

Materia, S., Borrelli, A., Bellucci, A., Alessandri, A., Di Pietro, P., Athanasiadis, P., . . . Gualdi, S. (2014). Impact of atmosphere and land surface initial conditions on seasonal forecasts of global surface temperature. *Journal of Climate, 27*(24), 9253-9271.

Miralles, D., Van Den Berg, M., Teuling, A., & De Jeu, R. (2012). Soil moisture-temperature coupling: A multiscale observational analysis. *Geophysical Research Letters, 39*(21).

Miralles, D. G., Teuling, A. J., Van Heerwaarden, C. C., & de Arellano, J. V.-G. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience, 7*(5), 345.

Miyoshi, T., & Yamane, S. (2007). Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Monthly Weather Review, 135*(11), 3841-3861.

Orsolini, Y., Senan, R., Balsamo, G., Doblas-Reyes, F., Vitart, F., Weisheimer, A., . . . Benestad, R. (2013). Impact of snow initialization on sub-seasonal forecasts. *Climate dynamics, 41*(7-8), 1969-1982.

Orth, R., Dutra, E., & Pappenberger, F. (2016). Improving weather predictability by including land surface model parameter uncertainty. *Monthly Weather Review, 144*(4), 1551-1569.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., . . . Yorke, J. A. (2004). A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography, 56*(5), 415-428.

Ott, E., Patil, D., Kalnay, E., Corazza, M., Szunyogh, I., Hunt, B., & Yorke, J. A. (2002). *Exploiting local low dimensionality of the atmospheric dynamics for efficient ensemble Kalman filtering*. Retrieved from

Paloscia, S., & Pampaloni, P. (1988). Microwave polarization index for monitoring vegetation growth. *IEEE Transactions on Geoscience and Remote Sensing, 26*(5), 617-621.

Pan, M., Cai, X., Chaney, N. W., Entekhabi, D., & Wood, E. F. (2016). An initial assessment of SMAP soil moisture retrievals using high-resolution model simulations and in situ observations. *Geophysical Research Letters, 43*(18), 9662-9668.

Perkins, S., Alexander, L., & Nairn, J. (2012). Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophysical Research Letters, 39*(20).

Piepmeier, J. R., Johnson, J. T., Mohammed, P. N., Bradley, D., Ruf, C., Aksoy, M., . . . Wong, M. (2014). Radio-frequency interference mitigation for the soil moisture active passive microwave radiometer. *IEEE Transactions on Geoscience and Remote Sensing, 52*(1), 761-775.

Prodhomme, C., Doblas-Reyes, F., Bellprat, O., & Dutra, E. (2016). Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe. *Climate dynamics, 47*(3-4), 919-935.

Quesada, B., Vautard, R., Yiou, P., Hirschi, M., & Seneviratne, S. I. (2012). Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change, 2*(10), 736-741.

Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources, 31*(11), 1411-1418.

Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water resources research, 44*(3).

Reichle, R. H., & Koster, R. D. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters, 31*(19).

Reichle, R. H., & Koster, R. D. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters, 32*(2).

Reichle, R. H., Koster, R. D., De Lannoy, G. J., Forman, B. A., Liu, Q., Mahanama, S. P., & Touré, A. J. J. o. c. (2011). Assessment and enhancement of MERRA land surface hydrology estimates. *24*(24), 6322-6338.

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P., Njoku, E. G., & Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research: Atmospheres, 112*(D9).

Reichle, R. H., McLaughlin, D. B., & Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review, 130*(1), 103-114.

Reichle, R. H., Walker, J. P., Koster, R. D., & Houser, P. R. (2002). Extended versus ensemble Kalman filtering for land data assimilation. *Journal of hydrometeorology, 3*(6), 728-740.

Rodell, M., & Houser, P. (2004). Updating a land surface model with MODIS-derived snow cover. *Journal of hydrometeorology, 5*(6), 1064-1075.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., . . . Bosilovich, M. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society, 85*(3), 381-394.

Schmugge, T., O'Neill, P. E., & Wang, J. R. (1986). Passive microwave soil moisture research. *IEEE Transactions on Geoscience and Remote Sensing*(1), 12-22.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., . . . Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews, 99*(3), 125-161.

Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., . . . Oleson, K.

W. (2006). Soil moisture memory in AGCM simulations: analysis of global land–atmosphere coupling experiment (GLACE) data. *Journal of hydrometeorology, 7*(5), 1090-1112.

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land–atmosphere coupling and climate change in Europe. *Nature, 443*(7108), 205-209.

Seo, E., Lee, M.-I., Jeong, J.-H., Koster, R. D., Schubert, S. D., Kim, H.-M., . . . Scaife, A. A. (2018). Impact of soil moisture initialization on boreal summer subseasonal forecasts: mid-latitude surface air temperature and heat wave events. *Climate dynamics*. doi:10.1007/s00382-018-4221-4

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate, 19*(13), 3088-3111.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D., Hunt, B. R., Kalnay, E., . . . Yorke, J. A. (2005). Assessing a local ensemble Kalman filter: Perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus A: Dynamic Meteorology and Oceanography, 57*(4), 528-545.

Teng, H., Branstator, G., Tawfik, A. B., & Callaghan, P. (2019). Circumglobal Response to Prescribed Soil Moisture over North America. *Journal of Climate*(2019).

Ushio, T., Okamoto, K. i., Iguchi, T., Takahashi, N., Iwanami, K., Aonashi, K., . . . Inoue, T. (2003). The global satellite mapping of precipitation (GSMaP) project. *Aqua (AMSR-E), 2004*.

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K. i., Aonashi, K., . . . Kachi, M. (2009). A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *Journal of the Meteorological Society of Japan. Ser. II, 87*, 137-151.

van den Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R. D., Seneviratne, S. I., & Camargo, H. (2012). Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate dynamics, 38*(1-2), 349-362.

Van Den Hurk, B., Jia, L., Jacobs, C., Menenti, M., & Li, Z.-L. (2002). Assimilation of land surface temperature data from ATSR in an NWP environment--a case study. *International Journal of Remote Sensing, 23*(24), 5193-5209.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., . . . Schneider, S. (2013). The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift, 22*(1), 5-33.

Weisheimer, A., Doblas-Reyes, F. J., Jung, T., & Palmer, T. (2011). On the predictability of the extreme summer 2003 over Europe. *Geophysical Research Letters, 38*(5).

Xie, P., & Arkin, P. A. (1997). Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American*

Meteorological Society, 78(11), 2539-2558.

Zaitchik, B. F., Macalady, A. K., Bonneau, L. R., & Smith, R. B. (2006). Europe's 2003 heat wave: A satellite view of impacts and land–atmosphere feedbacks. *International Journal of Climatology, 26*(6), 743-769.

Acknowledgement

박사학위 논문을 정리하며 제가 공부하는 동안 도움을 주신 많은 분들에게 감사드리며, 짧게나마 감사인사를 여기에 적어보려고 합니다.

대학교 2 학년 일 때 처음 학부지도교수님으로 처음 만나 근 10 년의 시간동안 저를 가르쳐 주신 이명인 교수님께 진심으로 감사드립니다. 학부생 시절 먼저 연락을 주셔 학부 인턴의 기회를 주신 인연을 시작으로 선배가 없는 유니스트 1 기인 제가 객지에서 힘들어할 때 선배대신 친근하게 상담해주시고, 전공수업을 들을 때 궁금한 것을 가져가면 교수님과 함께 탁자에 앉아서 공부했던 기억이 많이 남습니다. 그런 인연이 대학원까지 이어져서 박사학위과정 동안 아낌없는 지원을 해주시고, 학문적으로나 인성적으로 부족한 저를 올바른 길로 지도해 주셔서 진심으로 감사드립니다. 앞으로도 대학원 시절의 기억을 평생 기억하면서 열심히 하도록 하겠습니다.

그리고 대학원 기간동안 허점이 많은 연구를 탄탄해지도록 많은 도움을 주신 박사학위 심사위원 교수님들께도 진심으로 감사하다는 말씀드립니다. 박사 최종발표를 준비하면서 제안발표 자료를 보니 허점투성이 였던 연구를 교수님들의 조언을 차례로 보완해보니 어느덧 박사학위의 연구로 정리할 수 있었던 것 같습니다. 광주에서 울산까지 반나절동안 이동해서 심사발표에 참석해주시고, 자료동화에 대해서 전무한 저에게 학위과정에 많은 도움을 주신 함유근 교수님 정말 감사드리고, 항상 복도에서 온화한 미소로 반갑게 인사를 건네며 마음을 편안하게 해주신 임정호 교수님 진심으로 감사합니다. 그리고 UM 모델이 한국에 도입된 초창기 기상대를 다니면서 힘들게 연구할 때 함께 열악한 환경에서 연구하면 정도 많이 들고, 제가 힘들 때 항상 웃으면서 학생입장에서 많이 격려해 주셨던 차동현 교수님, 대학원 초창기 연구 프로젝트를 하며 힘든 시기를 겪는 동안 많은 격려와 연구적으로 큰 도움을 주셔 학위과정에서 큰 도약을 하는데 도움을 주신 정지훈 교수님께도 진심 어린 감사의 말씀을 전합니다. 그리고 저의 박사학위 심사위원은 아니셨지만 학부생 인턴 시절부터 연구의 꼼꼼함이나 호기심에 대한 탐구를 할 수 있는 연구자세를 알려주신 김대현 교수님, 일본 해외학회때 알게 된 이후로 졸업할 때까지 대학원 생활동안 항상 대학원생의 고충을 이해해 주시며 큰 도움을 주신 윤진호 교수님, 바다 건너 멀리 계시지만 가끔 볼때마다 항상 응원해주시고 졸업 후에도 많은 연구기회를 주신 김형준 교수님께도 감사의 말씀을 전합니다.

또한, 대학원생활동안 연구과제를 하면서 기상청에 계시는 많은 분들에게 도움을 받아 기상청에 계시는 많은 분들에게도 감사의 말씀을 드리고 싶습니다. 대학원 초창기 GloSea5 모델을 배우기 위해서 1 주의 절반을 제주도에서 보냈는데 항상 귀찮은 내색없이 도와주신 강현석 박사님, 부경온 박사님, 이조한 연구관님 모두 정말 감사드립니다. 특히,

저를 좋게 봐주시고 대학원 중에 UK Met Office, Bureau of Meteorology 에도 가서 발표를 할 수 있게 해 주신 강현석 박사님께 다시한번 감사의 말씀드립니다. 덕분에 대학원 생활동안 좋은 동기를 가지고 연구를 할 수 있었던 것 같습니다. 그리고 저의 연구에 많은 관심을 갖고 함께 연구할 기회를 주시고 연구에 관심을 갖아 주신 김윤재 박사님, 임윤진 박사님, 현유경 박사님께 진심으로 감사드립니다. 박사님들의 관심 덕분에 항상 제주도 출장길을 향할 때 가벼운 발걸음으로 기분 좋은 출장이 되었던 것 같습니다. 앞서 연구하신 분들이 저에게 보여주셨던 좋은 모습, 좋은 말씀들을 깊이 새겨 항상 저도 후배들에게 좋은 본보기가 될수록 노력하도록 하겠습니다.

대학원 생활을 하는 동안 항상 가족같이 챙겨주고 희로애락을 함께한 연구실 멤버들에게도 정말 감사합니다. 우선 연구실 1 호 박사인 동민이형 형한테는 정말 고마운 것도 많고 그만큼 미안한 것도 많네요. 항상 부족함 많은 저에게 좋은 말도 많이 해주고 부족한 점도 많이 이야기해줬는데, 제가 선배가 되니까 그런 말을 해준다는 게 얼마나 고마운 것인지 늦게 알게 됐습니다. 성윤이형도 정말 저와 함께 고생 많이 하셔서 형이 미국에서 박사 하러 나간 이후로도 항상 힘들 때 마다 형 생각이 많이 들었고, 항상 든든한 고목처럼 항상 그 자리에 있어줘서 고맙습니다. 열심히 학위 마무리하시고 곧 좋은 소식 있길 바다건너에서 응원하겠습니다. 미래 김박사님! 그리고 가장 할말이 많은 대현이. 대학원 생활을 돌아보면 내가 행복해했던 기억에는 항상 너와 함께한 기억들이 대부분이더라, 아마 대학원 생활을 너와 함께 하지 않았다면 지금 내가 이자리에 있을지 모르겠다. 나보다 나이는 어리지만 나보다 어른 같고, 부족한 나를 많이 도와준 대현에게 다시한번 고마운 마음을 전합니다. 그리고 지금 졸업을 하고 나간 민상, 우리 연구실 1 호 외국인 학생인 Son 모두 각자 연구실 졸업 후 열심히 연구하고 있는 모습을 항상 응원하고 곧 좋은 소식이 있길 바랍니다. 지금 연구실에 있는 엄마 같은 혜림, 걱정이 많이 되는 강한, 후배들에게 인기가 많은 낙빈, 연구실 분위기 메이커 승희, 서글서글한 부산남자 선래, 웃음이 많은 지해, 그리고 연구실에 오신지 3 개월 남짓 됐는데 3 년정도 계신 것처럼 친화력이 좋으신 이준리 박사님 모두 제가 연구실 생활하는데 많은 도움 주셔서 감사드립니다.

그리고 제가 학교에서 즐겁게 학부, 대학원 생활까지 잘 마치게 해준 대학/대학원 동기들, 후배들에게 감사의 말씀을 전합니다. 함께 수학하면 동고동락하며 저의 넉두리를 들어준 학부 친구들 자칭 공주라는 선영, 성우, (정)승후, 민준, 승진과 추억 쌓기를 좋아하는 저와 함께 좋은 추억을 만들어준 상균, (한)대현, 철희, 동진, 재세 에게도 고마운 마음을 보냅니다. 여러분들과 함께한 설악산 설산 등반은 잊지 못할 듯합니다. 그리고 저와 짧은 시간이지만 룸메를 하면서 많이 친해진 철희, 기대반 걱정반의 마음으로 학교를 떠나며 멀리서도 항상 응원을 약속합니다. 재홍이와 함께 밤새 싶은 대화를 나눈 날들이 너무 생각날 것 같고, 늦게 친해져 함께 한 시간이 짧아 아쉬운

마음이 남습니다. 그리도 대학원 대부분의 기간동안 룸메를 하며 외롭지 않게 말동무 친구가 되어준 홍대, 학부생 시절부터 룸메하며 인연을 지속한 해원, 규태, 현수, 정 모두 고맙습니다. 그리고 대학원생활동안 진성 축구인과 함께 공을 차느라 고생한 민수, 진영, 보광, 민호, 재원, 상준, 재범, 시호, (정)승후 모두 기억에 많이 남을 것 같습니다. 항상 집에 올라갔을 때 반겨준 서울에 있는 동네 친구 진, 중경, 도훈, 준섭, 균, 동우, 자원, 그리고 대학동기 다정이도 오래도록 감사하도록 하겠습니다.

20 대의 시간을 유니스트에서 보내며 정말 좋은 선후배들과 교수님들을 만나 즐거운 추억들을 많이 쌓고 많은 것을 배웠습니다. 앞으로도 말을 많이 하기보다는 많이 듣고 공감해줄 수 있는 사려 깊은 사람이 되도록 노력하겠습니다.

마지막으로 저의 학위 과정을 묵묵히 기다리며 응원해준 혜원에게 진심으로 고마움을 전하고, 무엇을 하던지 진심으로 지원해주시고 격려해주신 부모님과 누나 정말 너무 고맙고 사랑합니다. 학위 하는 동안 가장 가까운 사람들이 항상 그 자리에 있어주는 것만으로 가장 큰 행운이라고 하는데, 저 에게도 그런 큰 행운아래에서 공부할 수 있게 해 주셔서 진심으로 감사합니다. 학위를 하면 모든 것이 끝날 줄 알았는데, 이제서야 비를 가려주는 지붕에서 벗어서 새로운 인생의 출발점에서 새롭게 서있는 것 같습니다. 다가올 시간에 쉽지 않은 도전과 고난이 많겠지 앞서 감사드린 여러분들과의 기억을 추억하면서 현실에 타협하지 않고 제 소신을 굽히지 않으며 즐겁게 연구하며 살겠습니다.